Stress responses and sugar metabolism in *Bacillus subtilis*: a transcriptomic portrait

Andrzej Lulko

Marcie i dzieciom ...

The work described in this thesis was carried out in the Molecular Genetics research group of the Groningen Biomolecular Sciences and Biotechnology Institute (Faculty of Mathematics and Natural Sciences, University of Groningen) and was financially supported by the grant IGE01018 of the Innovation Oriented Research Programme on Genomics (IOP Genomics) of the Dutch Ministry of Economic Affairs.

This thesis was printed by IpskampDrukkers with financial support of the Groningen Biomolecular Sciences and Biotechnology Institute (GBB).

ISBN 978-90-367-4693-9 (printed version)
ISBN 978-90-367-4694-6 (digital version)

Cover was designed by the author. The image of *B. subtilis* DNA microarray was provided by Siger Holsappel.

RIJKSUNIVERSITEIT GRONINGEN

Stress responses and sugar metabolism in *Bacillus subtilis:* a transcriptomic portrait

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
vrijdag 7 januari 2011
om 14.45 uur

door

Andrzej Tadeusz Lulko

geboren op 5 april 1977
te Wrocław, Polen

Promotor:	Prof. dr. O.P. Kuipers
Copromotor:	Dr. G. Buist
Beoordelingscommissie:	Prof. dr. T. Abee
	Prof. dr. J. M. van Dijl
	Prof. dr. M. Kleerebezem

Contents

Chapter 1

General introduction
and scope of the thesis

Bacillus subtilis, the model organism for Gram-positive bacteria

Bacillus subtilis is a Gram-positive, rod-shaped bacterium commonly found in soil. It displays several features, which render this organism an interesting object not only for the scientific community but also for industrial applications. These features include amongst others the capability to:

(i) take up exogenous DNA (facilitates genetic manipulations which are well-established for this organism);

(ii) form endspores (so called sporulation, offers a model for studying this relatively simple developmental process);

(iii) secrete large quantities of proteins (also including heterologous proteins).

In addition, it was the first Gram-positive bacterium for which the complete genome sequence became available in 1997 [137]. This revealed that the chromosomal DNA of *B. subtilis* 168 strain has a low G+C content (43.5%) and comprises around 4.1 million base pairs, corresponding to more than 4100 genes. Remarkably, at that time only 30% of these genes had a well-known function, while for approximately half of the remaining ones an annotated function could be assigned based on sequence homology. Still, more than 1400 open reading frames (ORFs) were left without even a putative function assigned, being an enormous challenge for the scientific community. Therefore a functionial analysis programme (*B. subtilis* Functional Analysis, BSFA) was initiated, in which a collection of integrational reporter mutants of most of unknown genes (so-called "y-genes") was generated and phenotypically characterized. The BSFA programme revealed that some of these genes are required for stress resistance or are involved in the competence development or sporulation processes [90]. Due to the wealth of information available with respect to physiology and functional genetics, *B. subtilis* became a perfect model organism for Gram-positive bacteria. Moreover, the availability of the genome sequence enabled the development of transcriptomics technologies (described in more detail in chapter 2) that allow simultaneous monitoring of expression levels of almost all genes. These sophisticated methods offered an unprecedented opportunity to enter unexplored territories of *B. subtilis* research.

Regulation of transcription in *B. subtilis*

With more than 4100 genes in its genome, it is a challenge for the *B. subtilis* cell to manage its genetical material within a small space confined by the cellular envelope. Due to energy consumption reasons, only a fraction of the total number of genes is constitutively expressed (so-called housekeeping genes) [205]. The products of these genes are crucial for the cell at any circumstances (for example enzymes of the glycolytic pathway or ribosomal proteins). On the other hand, many other proteins are only needed "upon request" in specific situations, such as the availability/lack of certain compounds or when being under stress conditions. To recognize environmental changes and adequately respond to them, bacteria possess regulatory molecules, so-called transcription factors (TFs) that control expression of

one or more genes [254]. The process of gene expression is initiated by binding of the RNA polymerase (RNAP)-sigma factor complex to the promoter/operator site (DNA sequence in front of a gene) that is recognized by the sigma factor component [254]. TFs bind to the promoter sequence as well and may either facilitate or hamper the attachment of the RNAP leading to activation or repression of transcription, respectively and as such they can be considered a key element in regulation of gene transcription. Although there are more than 250 predicted TFs and 17 sigma factors encoded in the genome sequence of *B. subtilis*, only a limited number of them have a defined function [81,226]. In general, global transcription regulators and specific regulators can be distinguished based on the number of regulated genes in their regulons, the former driving the expression of a large number of genes of different functional categories, whereas the latter regulate only a small number of genes in a particular (metabolic) pathway.

Several factors contribute to the complexity of prokaryotic gene regulation by TFs:

(i) global regulators have the capability to act as either repressors or activators or both,

(ii) a gene can be regulated by more than one TF and

(iii) one TF can modify expression of another TF.

Moreover, external signals sensed by the cell are translated into a specific profile of intracellular metabolic cofactors that can act as signalling molecules capable of binding to TFs and thereby modulate their activities [254]. The interplay between three global regulators and their metabolites in *B. subtilis*, i.e. CcpA, CodY and TnrA, perfectly exemplifies this complexity and illustrates how gene regulatory systems intertwine in order to obtain the most efficient exploitation of available carbon and nitrogen sources [228]. Last but not least, a plethora of other regulation mechanisms adds up to the complexity of bacterial transcriptional regulation. Especially the field of posttranscriptional regulation by regulatory RNAs, including riboswitches, non-coding RNAs (nc-RNAs) and antisense RNAs, have gained a lot of interest in prokaryotic research [262]. The development of technologies such as tiling microarrays and RNA-seq (deep-sequencing) is expected to revolutionize the manner in which transcriptomes will be analyzed in the near future ([229,259], see also chapter 7). A prominent example of the implementation of the tiling arrays is the pioneering work of Rasmussen and colleagues, who identified 84 putative nc-RNAs and 127 antisense transcripts in *B. subtilis* [205].

Databases such as PRODORIC or DBTBS (a database of transcriptional regulation in *B. subtilis*) aim at constructing regulatory networks that depict relationships between TFs, promoter sequences as well as operon and regulon organization based on literature information [81,226]. The latter database compiles information on 35% of all *B. subtilis* promoters and offers valuable help with the interpretation of the enormous amount of data derived from whole-genome analysis approaches, such as transcriptomics or proteomics.

Carbon catabolite regulation and metabolic pathways under control of CcpA

Bacteria need to manage their energy sources in the most efficient manner to achieve the fastest possible growth rate under particular circumstances, as well as to survive in a highly competitive environment. Carbon catabolite regulation is a mechanism that allows cells to utilize the available carbon sources in a preferential order [235]. For example, in the presence of glucose, the expression of genes for metabolism of other, less preferred carbon substrates is switched off. The CcpA protein [98,106,174] is a master regulator of carbon catabolite control in many low-GC Gram-positive bacteria [261]. This regulator is a member of the LacI/GalR family of transcriptional regulators [265], which can act as a repressor in carbon catabolite repression (CCR) and as an activator in carbon catabolite activation (CCA) [104,235].

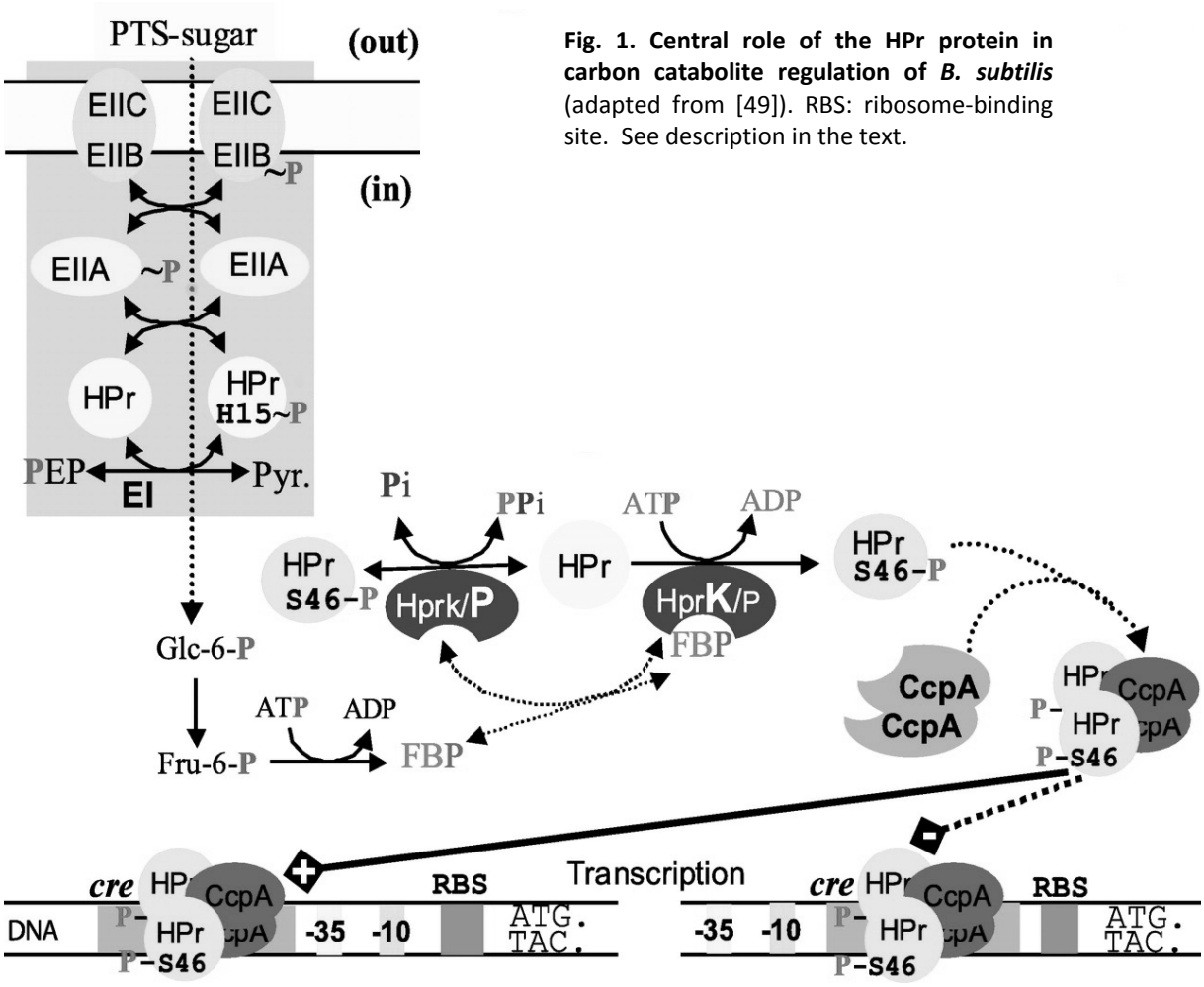


Fig. 1. Central role of the HPr protein in carbon catabolite regulation of *B. subtilis* (adapted from [49]). RBS: ribosome-binding site. See description in the text.

Independent of the availability of preferred carbohydrates in the medium, CcpA is constitutively expressed [174], which implies that this regulator needs additional factors to exert its pleiotropic function. CcpA forms a dimer [219] and its DNA-binding activity is stimulated by complex formation with HPr-Ser-P [52,75] or HPr-like protein Crh-Ser-P [77].

In addition to these phosphoproteins also low molecular weight molecules such as NADP, glucose-6-phosphate (G6P) and fructose-1,6-bisphosphate (FBP) can modulate either DNA-binding properties or the interaction with the transcriptional machinery of the CcpA-(HPr-Ser-P) complex [79,125,126,223]. The phosphoenolpyruvate:sugar phosphotransferase system (PTS) is a multiprotein system that allows bacteria to transport hydrophilic compounds (such as sugars) through the membrane. Next to its role in transport it has also a function in signal transduction. It consists of three general components: enzyme I (EI), Hpr protein and several sugar-specific types of enzymes II (EII) consisting of three or four domains (EIIA, B, C, (D)). The phosphoryl residue is transferred from PEP via EI, the HPr protein and enzymes IIA and IIB to the sugar incoming via the membrane-bound IIC (Fig. 1) [235]. The HPr protein acts as a central sensor of carbon metabolism in *Bacillus* species, as it can be phosphorylated by EI of the PTS or by HPr kinase/phosphorylase (HPrK/P) at two distinct positions, either His-15 or Ser-46, which takes places under low or high intracellular glucose concentrations, respectively [51,235]. HPr-His-P is involved in PTS sugar uptake from the environment (out) into the cell (in) and regulation of catabolic enzymes or regulators with a PTS-regulation domain (PRD) [197,235]. Oppositely, HPr-Ser-P is a predominant form of the HPr protein under high cellular glycolytic activity [53,196], leading to high intracellular concentrations of 1, 6-bisphosphate (FBP) and ATP, and this form serves as a cofactor to significantly enhance CcpA binding to target sites of various catabolic genes and operons [75,115,223] causing the repression or activation of their transcription (Fig. 1).

CcpA binds a DNA regulatory sequence known as catabolite-responsive element (*cre*) of which the consensus sequence TGWAANCGNTNWCA (N=any base; W=A or T) was first detected in the promoter region of *amyE* [105,112,173,266]. Later it was discovered that the presence of *cre* either upstream or in the promoter region, or inside a gene does not necessarily actual imply regulation by CcpA [173]. On basis of *in vivo* operational *cre* sites the following consensus sequence, WWTGNAARCGNWWWCAWW, was proposed. This consensus suggests a high level of degeneracy and indeed, depending on the search method and consensus sequence used, different numbers of *cre* sites can be predicted in the *B. subtilis* genome [50]. The position of the *cre* site(s) relative to the transcription start (TS) of a transcriptional unit determines the regulatory effects upon HPr-Ser-P/CcpA complex binding. In general, promoters with a *cre* site present upstream of the hexameric -35 sequence undergo transcriptional activation as in case of *ackA* and *pta* [198,247]. Interestingly, both genes contain an additional conserved sequence upstream of the CcpA binding site, the presence of which appears to be crucial for transcriptional activation of *ackA* [176]. The presence of a possible upstream activating region (UAR) has been also documented for other CcpA-activated genes (see chapter 5).

Recently, a direct positive regulation by CcpA was also demonstrated for the *ilvB* promoter [225,244]. Transcriptional repression by binding of CcpA to a *cre* site downstream of the TS, which blocks elongation by RNAP, has been shown. This transcriptional roadblock mechanism has been proposed for the *xyl*, *ara* and *gnt* operons [110,112,281] and for *sigL* [37]

and *acsA* [282]. In the latter case, affinity of CcpA for the *cre* site was also dependent on the composition of the nucleotides flanking *cre*. In chapter 5 the *ydhMNOPQRST* operon and the *ycgN* gene are identified as novel CcpA-repressed targets, for which putative *cre* sites were detected upstream of the transcription start site and within the ORF, respectively.

Prevention of binding of RNAP to the promoter sequence has been shown for the *acuABC* and *bglPH* operons where a *cre* site overlaps with parts of the promoter region, which may prevent transcription initiation [82,135]. It has also been suggested that CcpA does not prevent RNAP from binding to the promoter of *amyE* but that it interacts directly with the RNAP complex already bound to its operator site [127]. In addition, activation or repression by CcpA binding to *cre* is helix-face dependent since non-integral turns of helix insertions caused relief of CCR of *amyE* [127] or lack of activation as in case of *ackA* [247]. Novel activated members of the CcpA regulon, including *opuE* and the *opuAABC*, *yhb* and *man* operons, which all have a putative *cre* site that appears to be dependent on helical topology, are the result of the microarray analyses presented in chapter 5.

CcpA, as a global regulator, directly or indirectly regulates expression of several hundreds of genes during the growth cycle of *B. subtilis* as has been shown by transcriptome analysis at different growth phases [154]. Pleiotropicity of *ccpA* knock-out mutants has also been shown by various transcriptomics studies performed at mid-exponential phase of growth [18,150,181,276]. Moreover, it has also been demonstrated that besides its role in carbon metabolism, CcpA also regulates phosphorus metabolism [36,201] as well as nitrogen metabolism [37,65,258], especially via its involvement in regulation of the genes required for the biosynthesis of branched-chain amino acids (BCAA) [152,225,244,290]. BCAAs are very important intracellular metabolites as they serve not only as signalling molecules but also as major building blocks of proteins. In addition these compounds are precursors for the fatty acids of the membrane phospholipids. The relevance of BCAAs is reflected by the fact that their concentration in the cells is regulated by three global regulators - CodY, TnrA and CcpA that respond to the levels of GTP/BCAAs, glutamine and FBP, respectively [228]. Sophisticated and fine-tuned regulation of the *ilvB* operon for biosynthesis of branched-chain amino acids in *B. subtilis* is an excellent example of the interplay of these three transcriptional factors [225].

The crystal structures of the ternary complexes of (P-Ser-HPr)-CcpA-DNA and (P-Ser-Crh-HPr)-CcpA-DNA are available for *Bacillus megaterium* [219,220]. These structures revealed that CcpA is composed of an N-terminal DNA binding domain (residues 1-60) and a C-terminal domain that contributes to dimerization and corepressor (HPr or Crh) binding. The former domain binds to the DNA sequence through the helix-turn-helix (HTH) motif and the hinge helix, which also connects to the C-terminal domain. Upon HPr binding, N-terminal subdomains rotate, leading to the juxtaposition of the DNA-binding regions of the CcpA dimer that results in “hinge” helix formation [219]. This so-called allosteric switch mechanism in the presence of cognate DNA is a central event in P-Ser-HPr-dependent CCR. Fairly recently, several variants of CcpA containing single amino acid exchanges located in

different regions of CcpA of *B. subtilis* were investigated with regard to their transcriptional activities and it was shown that a variety of complex regulatory effects, which were dependent on the promoter as well as the type of the mutation, occurred [231].

Stress response in *B. subtilis*

Stress response – general considerations

In natural environments as well as during industrial processes, microorganisms are often challenged with a variety of unfavourable conditions. To increase their chance of survival, bacteria evolved various mechanisms to confront and adapt to encountered stresses and to maintain proper cellular physiology. In general, they achieve that by a concerted reprogramming of gene expression, leading to a synthesis of proteins that are utilized for a rapid and optimal adjustment to a particular stress.

The response of *B. subtilis* to diverse stress conditions, such as salt, oxidative, acid, alkaline, ethanol, or heat and cold shock stress has been thoroughly investigated and the major role of alternative sigma factor SigB in governing induction of many general stress-responsive genes became apparent [101,221]. *B. subtilis* and other Gram-positive bacteria have also developed other mechanisms to defend themselves against growth-restricting conditions by forming a complex adaptation network. The constituents of this defence system are elegantly controlled by various transcriptional regulators (Fig. 2). As a result, the synthesis of antibiotics or non-specific or specific stress proteins, including degradative enzymes and membrane transporters, is initiated and alternatively the processes such as chemotaxis, competence and finally sporulation can be developed [183]. Such a concerted action is directed in order to first neutralize the effects of the encountered stress and thereafter adapt to it and finally repair the damage [93]. Sporulation is the last resort of survival strategy of *B. subtilis* triggered by nutrient limitation and only a subpopulation of cells is capable of entering this complex developmental process orchestrated by five alternative sigma factors and the SpoOA global regulator. The remaining cell population has to employ an alternative strategy to starvation via the stringent response, in which the (p)ppGpp synthetase RelA plays a role in nutritional stress activation of SigB [285]. The dual action of RelA and SigB leads to so called “vegetative dormancy” characterized by reduced anabolic reactions (reduction of ribosome synthesis, DNA synthesis, cell wall synthesis, etc.) and a multiple preventive stress resistance. Besides, under some conditions, including high osmolarity and anaerobiosis, sporulation is not allowed, thus the cells have to rely only on alternative survival strategies [92]. Another survival strategy that *B. subtilis* employs as an adaptive response under unfavourable environmental conditions is motility and chemotaxis, both of which are governed by the alternative sigma factor sigD [96,164].

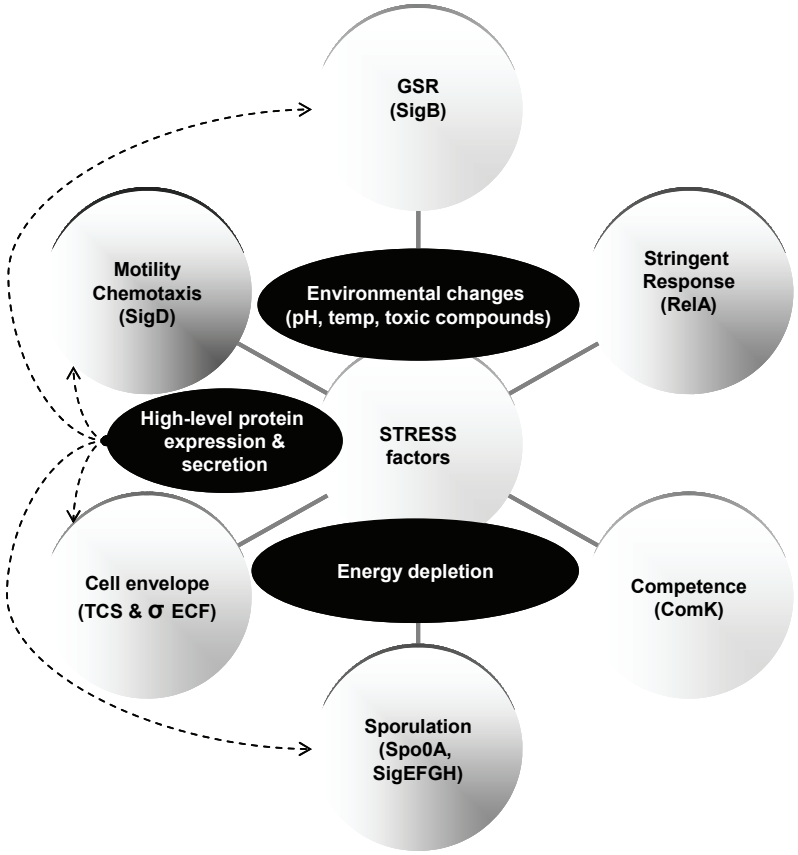


Fig. 2. General view on processes triggered by stress in *B. subtilis*. See text for details.

Cell envelope and secretion stress

At an initial stage, various stress factors pose a threat to the cell envelope. Therefore, maintaining its integrity is of crucial importance as it constitutes a first line of defence. It has been shown that complex regulatory mechanisms that consist of several two component signal transduction systems (TCS, see further in the text) and sigma factors of extracytoplasmic function (ECF) orchestrate responses to cell envelope stress.

Due to the fact that envelope stress occurs outside the cytoplasm, both systems consist of a membrane-anchored sensory component (sensor histidine kinase or anti-sigma factor) and a cytoplasmic transcriptional regulator (response regulator and ECF sigma factor, σ ECF). The sensory component receives a stress signal from the outside of the cytoplasmic membrane and induces the respective regulators, which in turn direct the expression of their target genes (Fig. 3). In case of TCS, activation of the response regulator is based on the transfer of a phosphoryl group from the histidine kinase. On the other hand, an anti-sigma factor inhibits its cognate sigma factor by direct protein-protein binding. Upon receiving a stimulus from the environment, the sigma factor is released, either by a conformational change or by proteolysis of the anti-sigma factor, and can bind to RNA polymerase to stimulate transcription. The anti-sigma:sigma pairs are analogous to the TCS but differ in the mechanism by which the sensor and regulator proteins communicate with each other

[95,116]. The physiological roles of several ECFs have also been investigated in relation to their contribution to other stress responses. SigX, similarly to the TCSs mentioned above, also controls genes related to cell wall metabolism and modification and provides protection against antimicrobial peptides [31] and contributes to survival of *B. subtilis* at high temperatures [102]. It was shown that the SigX regulon partially overlaps with the SigW regulon [103]. The latter one is induced by conditions affecting the cell envelope, such as cell wall antibiotics or alkaline shock [166]. Another important ECF in relation to stress is SigM, which is activated in response to cell wall antibiotics, ethanol, heat, acid, and superoxide [242].

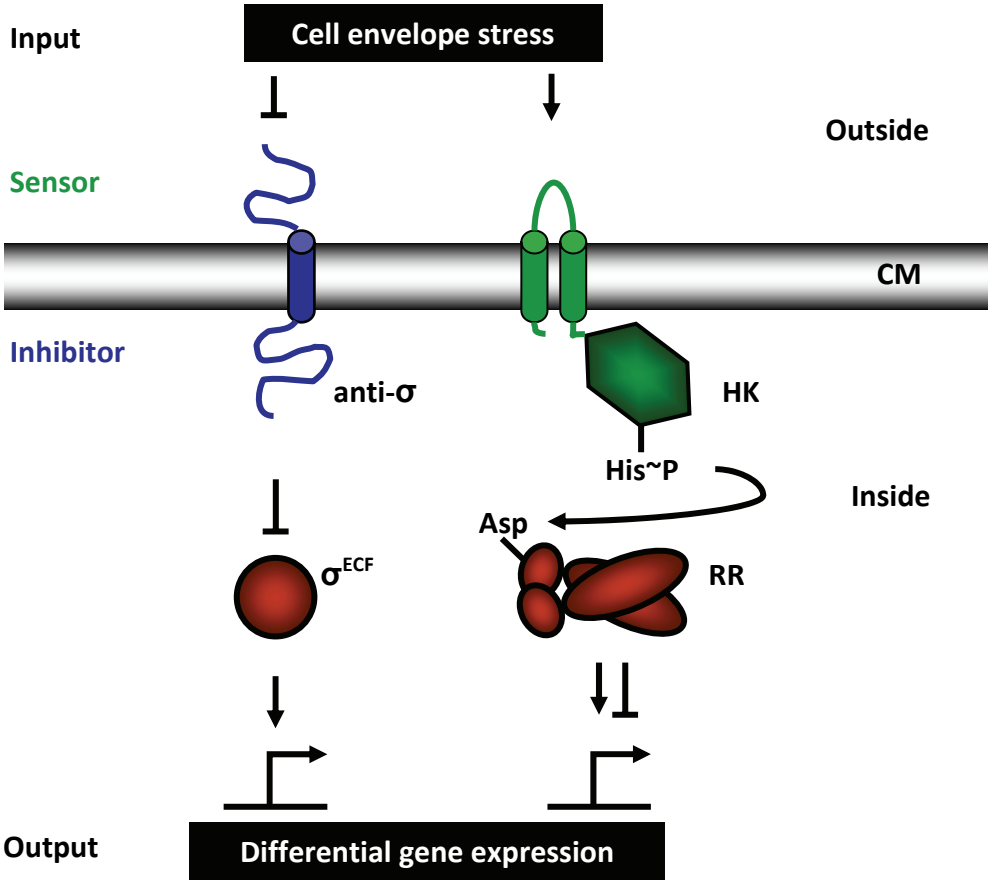


Fig. 3. Regulation of cell envelope stress by TCS and ECF sigma factors (adapted from [116]). Sensor proteins are shown in green, inhibitor in blue and transcriptional regulators in red. Arrows indicate activation, T-shaped lines repression. HK-histidine kinase; RR-response regulator; CM-cytoplasmic membrane. See text for details.

B. subtilis is capable of secreting high amounts of endogenous proteins into the extracellular medium [268]. Therefore, this bacterium and its relatives are often exploited as hosts for the production and secretion of heterologous industrially interesting enzymes. However, secretion of heterologous proteins in large quantities has been shown to lead to protein misfolding and their subsequent degradation [216]. In *B. subtilis*, accumulation of misfolded

proteins at the membrane-cell wall interface is sensed by the CssRS TCS, which consists of the membrane-embedded sensor kinase CssS and the response regulator CssR [108]. This system responds to high-level protein secretion and heat stress by phosphorylation of CssR which, in turn, activates transcription of the monocistronic *htrA* and *htrB* genes [48]. The CssRS system has also been shown to regulate the expression of its own operon [48,129]. HtrA and HtrB are membrane-bound serine proteases whose major functions are degradation of misfolded and aggregated proteins. Expression of both proteases is induced upon excessive expression of secreted proteins or a temperature increase [187]. Thus, the CssRS regulon forms a quality control and defence system when cells are confronted with secretion stress. The members of the CssRS quality control system become involved at that stage and take action in two possible manners, either by acting as a chaperone aiding refolding or preventing unfolding, or by HtrA/B-dependent proteolysis.

There are at least two other very important TCSs, which play an important role in monitoring cell envelope integrity and homeostasis, namely LiaRS and YycFG [17,167]. The former one has been shown to respond to a variety of stress factors, such as presence of external cell-wall antibiotics, alkaline shock, ethanol, organic solvents and secretion stress [116]. Although the latter one is *per se* not directly implicated in counteracting cell envelope stress, it is essential for viability and regulates expression of genes in cell wall metabolism under normal growth conditions [17]. Another TCS is DesRK, which is not directly involved in the cell wall metabolism, but has been shown to be essential under cold-shock conditions due to its role in controlling fluidity of the membrane [2,46,263].

Weak organic acid stress

The ability of weak organic acids to inhibit bacterial growth explains their widespread use in food production as effective antimicrobials. Although the mechanisms behind the inhibitory effects are not entirely elucidated, antimicrobial activity is assigned to the undissociated form of these compounds that can easily pass through the cell membrane [27]. Upon entry into the cytosol (close to neutral pH), the weak acid molecules dissociate into protons and anions, both of potential danger to the cell [42]. The intracellular excess of protons interferes with the pH gradient (ΔpH) across the membrane, causing an impaired cellular bioenergetics, since ΔpH, in addition to the membrane potential (ΔΨ), contributes to the transmembrane proton motive force (PMF) in bacteria [191]. Uncontrolled cytoplasmic protonation dissipates the PMF and hampers functioning of proteins and nucleic acids, which are crucial for survival of the cell. Moreover, toxicity of organic acids may result from the intracellular accumulation of anions that have deleterious effects on the cell as they increase osmolarity of the cytoplasm and also may influence activity of metabolic enzymes [213]. To deal with low pH stress (thus not specifically caused by weak acids), bacteria have developed various mechanisms of defence (Fig 4):

(i) A first line of defence has been assigned to the proton pumps such as the F_0F_1-ATPases, or other cation transport ATPases (Fig. 4A), which remove protons from

the cytoplasm at the expense of ATP in order to regenerate ΔpH and keep the cytosolic pH stable [42]. Through proton extrusion, the cell can temporarily maintain pH homeostasis and gain additional time to employ other strategies to respond to the encountered stress conditions, provided that they persist for a longer period of time.

(ii) One of these strategies is a decarboxylation reaction which involves the exclusion of intracellular protons (Fig. 4A). An example of this kind of reaction operating in Gram-positive bacteria is the glutamate decarboxylation/transport system (GAD). In this case, extracellular glutamate is transported into the cytoplasm, where it is converted to γ-aminobutyrate (GABA) via a proton-consuming, decarboxylation reaction, followed by its extrusion from the cell [42].

(iii) Low pH environments may also induce the expression of several proteins involved in the protection or repair of macromolecules such as DNA and proteins (Fig. 4B), which enables proper functioning of the cell [93]. As there are many of these kinds of proteins, only a few examples will be shortly described in this section. The RecA protein is a multifunctional protein serving a recombinase functioning in recombinational DNA repair and the SOS response in bacteria [157]. The nucleotide excision repair protein UvrA (part of the nucleotide excision repair pathway encoded by the *uvr* operon) has been shown to be involved in the repair of acid-induced DNA damage and adaptation to low pH [88]. Chaperones such as DnaK or GroEL have an important function under various suboptimal conditions as they prevent aggregation of proteins and repair the ones that have been damaged or misfolded by various stresses [156]. In addition, regulation by proteolysis of abnormal proteins formed under unfavourable conditions is of crucial importance, with Clp ATPases or Htr-like proteases playing a pivotal role in bacterial stress responses (chapter 3, [114,177]).

(iv) The role of TCS and sigma factors (Fig. 4C) in response to cell envelope stress has been discussed in the previous section.

(v) Also the cell density of bacteria affects their acid resistance (Fig. 4D). Cell density and biofilm growth were shown to modulate acid adaptation in *S. mutans* via a quorum sensing system (ComCDE) essential for cell density-dependent induction of genetic competence. The results of this study showed that mutants defective in the *comC*, *-D*, or *-E* genes had a diminished acid tolerance and that addition of synthetic competence stimulating peptide (CSP) to the *comC* mutant restored the acid tolerance [145]. Besides, formation of biofilms seems to physically protect cells located in the deepest layers from the elevated levels of acidity [42].

(vi) Other alternative responses to acidic stress are the modification of the cell wall or alteration of the cytoplasmatic membrane composition (Fig. 4E) as demonstrated for *S. mutans* [73], *B. subtilis* (chapter 4) and other bacteria [72].

(vii) Another possible mechanism of weak acid protection includes synthesis of alkali (Fig. 4F), for example via the urease and arginine deaminase (ADI) pathways through which ammonia is generated that can bind protons in the cytoplasm [45].

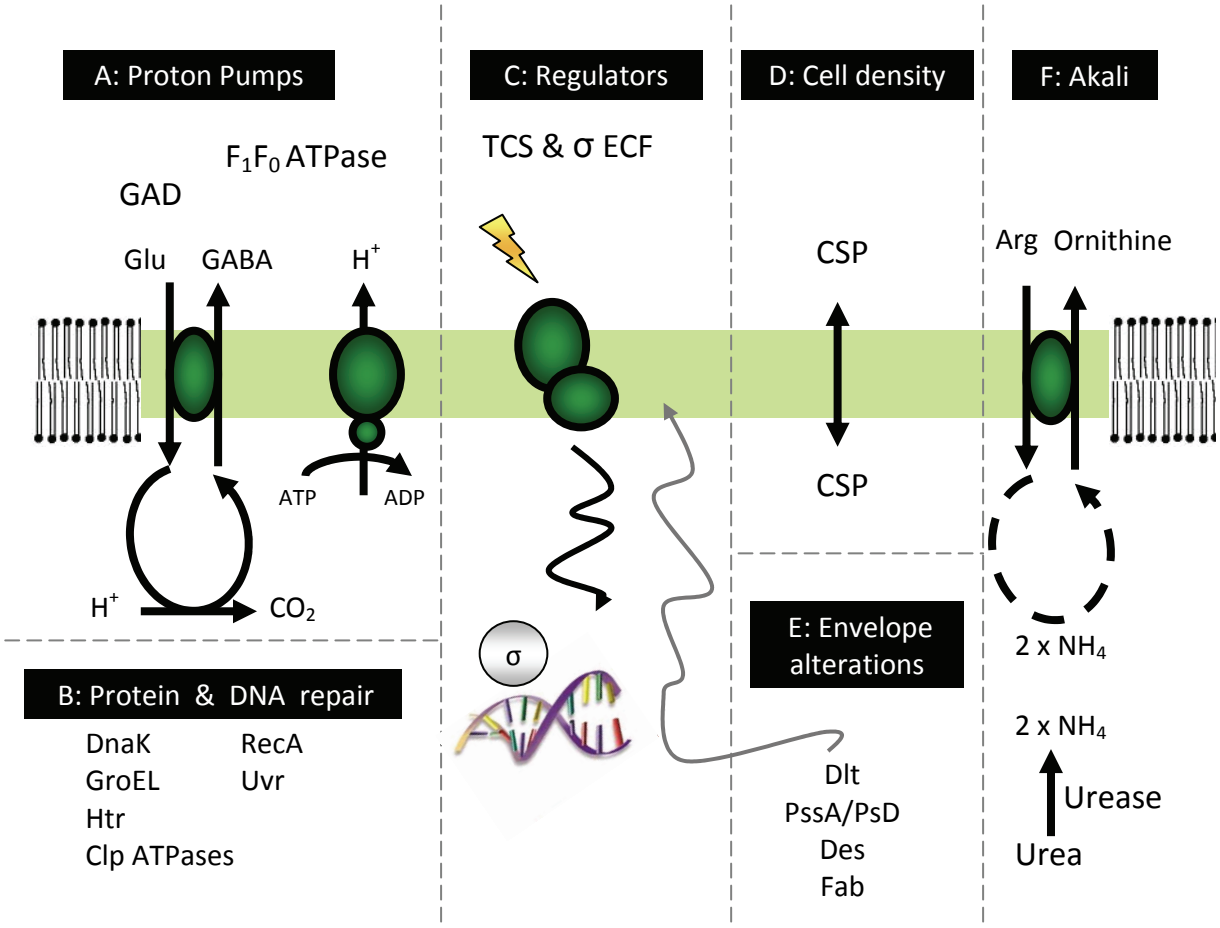


Fig. 4. Overview of the defence mechanisms against acid stress in Gram-positive bacteria (adapted from [42]). See text for details.

In conclusion, Gram-positive bacteria have a wide variety of different mechanisms at disposal to endure an (organic) acid challenge, though much remains to be established with respect to detailed analysis of this particular stress and its species-dependent specificity.
Only very recently a combined proteomics and transcriptomics approach, in which *B. subtilis* was treated with the salicylic acid, has started to tackle this question [57]. Remarkably, also in this case a major involvement of the SigB regulon was observed, which was exclusively attributed to salicylic acid, and not to the medium acidification alone. Furthermore, it was shown that salicylic acid and lactic acid in *B. subtilis* and *Lactobacillus plantarum*, respectively, activate the general stress response as well [57,194]. However, the role of this alternative sigma factor in relation to weak organic acids response is not as explicit as SigB, since it is not involved in response to sorbic acid [241].

The scope of this thesis

The work presented in this thesis was part of a project within the Innovation Oriented Research Programme (IOP) on Genomics titled "Comparative and predictive transcriptome analysis of Gram-positive bacteria for enhanced food functionality, quality and safety". There were two major goals of the project: (i) the development of the transcriptomics technology of four fully-sequenced and evolutionary related Gram-positive bacteria *B. subtilis*, *L, lactis*, *L. plantarum* and *B. cereus* and (ii) and the implementation of this technology to investigate the effects of various stress conditions encountered for example under industrial food fermentation conditions. The prerequisite of the project was the development of DNA microarray technology, which as described in **chapter 2** was a demanding and time-consuming challenge. This research focuses on the *B. subtilis* part of the IOP project and in general the effects of three diverse key stress-related topics are tackled: (i) overproduction of a heterologous protein (chapter 3), (ii) exposure to a weak organic acid (chapter 4) and (iii) disturbance of carbon-metabolism (chapter 5 and 6).

In **chapter 3,** a wild-type and a *cssS* mutant strains of *B. subtilis* are challenged with a secretion stress elicited by overproduction of a heterologous α-amylase at the late exponential and stationary phase of growth. Putative novel members of the CssRS regulon are identified and verified. It is shown that the genes involved in the processes of sporulation and motility are oppositely affected. This chapter is the only one in which the commercial nylon **macroarrays** and the method of radioactive probe labelling are used, since at that moment of time the glass microarray technology was not fully available yet.

In **chapter 4**, growth and global gene expression patterns in response to lactic acid exposure at different pH is analysed. Various mechanisms of adaptation to a mild lactic acid stress (pH 5.5) are observed. In addition, overexpression strains are identified and investigated to find potential candidates with improved growth characteristic under severe lactic acid challenge (pH 5.0).

In **chapter 5,** the regulon of the global transcriptional regulator of carbon metabolism is characterized at four different stages of growth. The approach allowed to dissect the CcpA regulon over time and permitted identification of new (putative) members of the CcpA regulon.

In **chapter 6**, the effects of CcpA point mutations on the activation or repression function of CcpA is evaluated. Four independent mutants containing single amino acid substitutions localized in the DNA-binding domain or the HPr(CrhP)-SerP binding domain of CcpA are investigated. Depending on the localization of a particular substitution, the transcriptome profiles resemble either the full knockout (indicating a big impact of the single point mutation) or the wild-type CcpA.

Finally, in **chapter 7** a global view of the results of the previous chapters and an outlook for future research are given. Furthermore, the author presents a short comparative analysis of the data obtained within the IOP project.

Chapter 2

DNA microarrays: experimental issues and their application in *Bacillus subtilis* gene expression profiling

Andrzej Lulko, Siger Holsappel, Jolanda van Munster, Girbe Buist and Oscar Kuipers

ABSTRACT

DNA microarray technology has been developed rapidly in the last decade and gained enormous popularity in many areas of microbial research. Due to the fact that this technology allows measuring changes in expression levels of thousands of genes in a single experiment, it provides a powerful tool for exploring the regulation of gene expression in bacteria. In this microreview special attention is paid to the basic principles of DNA microarray analysis and to the steps involved in the experimental workflow with emphasis on the authors' experience on fabrication of amplicon-based microarrays. Furthermore, a literature overview on the application of transcriptomics in *B. subtilis* research will be presented.

Introduction

DNA microarray technology

A DNA microarray (also called DNA chip) harbours a set of spots containing DNA fragments dried from solution, representing individual gene sequences usually from one organism. In theory, microarrays allow to determine the expression levels of all mRNA molecules (targets) present in a collection of cells. These mRNAs are encoded by genes active under given conditions of growth and will hybridize to their complementary DNA strands (probes) represented on the chip. Such microarrays can be considered as a multiple Northern blot experiment, since the same underlying principle is used, with the huge advantage of the former one that it allows including thousands of DNA probes on the surface of the chip. However, in contrast to Northern blots, microarrays cannot provide information on the length of the mRNA. First reports on the use of DNA microarrays for gene expression profiling appeared in 1995 and a complete eukaryotic genome (*Saccharomyces cerevisiae*) on a microarray was published in 1997 [140],[217].

Generally, a microarray experiment can be divided into several steps: microarray fabrication, RNA isolation, probe preparation and labeling, hybridization and data analysis, and each of these stages will be discussed further in the text. In this thesis special attention will be paid to the microarray production issues of *B. subtilis*.

DNA microarray platforms

DNA probes can be synthesized by use of PCR methods (amplicons) or by chemical synthesis of oligonucleotides (20 to 80 base pairs). Subsequently, the DNA is delivered either on coated glass microscope slides or on plastic or nylon membranes, by robotic systems to obtain oligoarrays or cDNA microarrays. Another possibility is to synthesize DNA probes directly (*in situ*) on the surface of the DNA chip (oligonucleotide gene chips) by implementing either photolithography (Affymetrix, 20-25 bp) or an ink-jet approach (Agilent, 50-80 bp). The array technology is constantly improving, a prominent example of which is the development of the tiling DNA microarrays. These kinds of arrays not only cover both strands of the genome (thus not only one strand as in the case of "traditional" microarrays) but also include intergenic regions of a genome [80,205]. The benefits of genomic tiling arrays will be more thoroughly discussed in chapter 7. The Affimetrix technology is rather expensive, inflexible and does not allow for co-hybridizations (one colour arrays, no ratios can be generated), but on the other hand it offers a high level of reproducibility and robustness. Oppositely, spotted microarrays are relatively cheap and can be designed according to the need of specific investigations. However, their production is burdened with many technical challenges that may lead to a high variability of in-house prepared slides. Some examples of these problems based on the author's experience during development of *B. subtilis* microarrays are given in the next section.

Fabrication of DNA-microarrays
(based on experience with the PCR-based microarrays of *B. subtilis*)

The first goal of DNA microarray fabrication is to obtain sufficient amounts of probe DNA, which is to be deposited on a suitable surface. In case of *B. subtilis* in 2001 it was chosen by the Molecular Genetics department in Haren to obtain PCR products for all of ORFs in its genome by using primers from Eurogentec (www.eurogentec.com). Each primer contained the same 15-base tag-sequence at the 5'-end followed by 18 to 22 bases of ORF-specific sequence. The first round of PCR reaction (primary PCR) was done by Eurogentec by using genomic DNA of *B. subtilis* 168 as template and the specific part of the oligo. The second round of PCR reactions (secondary PCR) was performed at the Molecular Genetics group by using diluted primary PCR products as templates and the C-terminal amino-modified universal primers complementary to the tag-sequence from the first PCR round. The second round PCR reactions were meant to (i) reduce chromosomal DNA contamination, (ii) increase yield and (iii) to introduce the amino-tag for a better spotting chemistry. The entire primary and the secondary PCR reactions were performed in 96 well-plates in 100 μl volumes and were controlled on standard gel electrophoresis for quality yield and for the expected size at Eurogentec or the Molecular Genetics group, respectively. Surprisingly, this task not only appeared to be more troublesome than initially anticipated, but also various bottlenecks were encountered, especially in the case of secondary PCR. In the first attempt, more than 800 PCR reactions of the second PCR round failed to yield correct amplicons, which corresponded to approximately 20% of all amplicons. As presented in Fig. 1, in many cases little or no PCR products as well as smears or multiple-band products were obtained. Different approaches were applied to these four groups, i.e. various PCR premixes or PCR conditions to obtain a proper PCR product. After four subsequent rounds of PCR attempts, there were still 223 amplicons for which PCR reactions failed. Most of these were either amplicons with missing (119 amplicons) or smeared (81 amplicons) PCR products. The sequences of these problematic amplicons were subjected to selection of the most unique sequences (not longer than 800 bp) within their ORFs and then primers were prepared based on these unique sequences by using the UniFrag and the GenomePrimer software, respectively.

UniFrag and GenomePrimer allow an automatic selection of unique regions within a DNA sequence and a design of primers in these regions, respectively, and have been successfully used for amplicon design and production of bacterial DNA-microarrays [252]. This method proved to be very effective and it delivered not only a high success rate but also relatively high and comparable amounts of PCR products (Fig. 1C).

Based on the experience gained during the amplification of the *B. subtilis* genome, instead of PCRing the whole ORFs, the same approach was also implemented later in time for the microarray design of *Lactococcus lactis* IL1403 and *Streptococcus pneumoniae* TIGR4 [136].

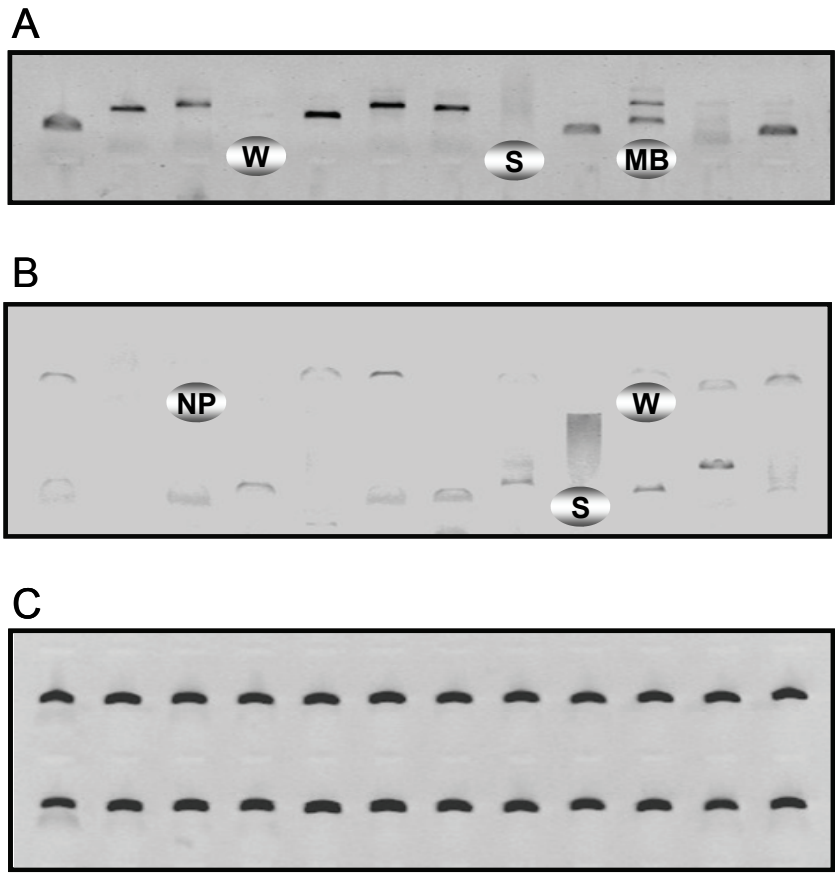


Fig. 1. Agarose gel separation of PCR fragments. Typical results of amplification of the whole (A, B) and short versions of ORFs (C) are given. W-weak product, NP-no product, MB-multiple-band product and S-smear.

A subsequent step of microarray preparation should involve purification of PCR products by ethanol precipitation or commercially available gel-filtration systems (Millipore or Qiagen for example) to remove unused primers, nucleotides, template DNA, proteins and other impurities present in the PCR mixture that in principle could influence the hybridization process. However, due to the laborious and time-consuming amplification progress and to reduce costs and possible loss of material, it was decided to skip the PCR purification step. This has been further supported by a publication of Diehl's *et al.,* in which evidence was presented that microarrays can also be manufactured from unpurified PCR products [55]. Although a test batch of *B. subtilis* DNA chips was successfully spotted and used in hybridization experiments, consecutive spotting events resulted in a deteriorated spot morphology (Fig. 2). After several rounds of spotting the PCR material became precipitated or degraded. Actually, the same problem occurred in more groups working with DNA microarrays (information from the gene-arrays list at itssrv1.ucsf.edu and http://tech.groups.yahoo.com/group/microarray/). As can be seen in Fig. 2 not only the overall signal intensity (for example see the middle arrow in Fig. 2.) but especially the spot morphology (upper and lower arrows in Fig. 2) was considerably deteriorated between the first (Fig 2A) and the third spotting run (Fig. 2B). A substantial number of spots gained a small shape in the form of a light central point (see the zoomed-in fragment in Fig. 2), which

resulted in loss of a great part of the array data as the signal of these kinds of "small" spots cannot be accurately quantified.

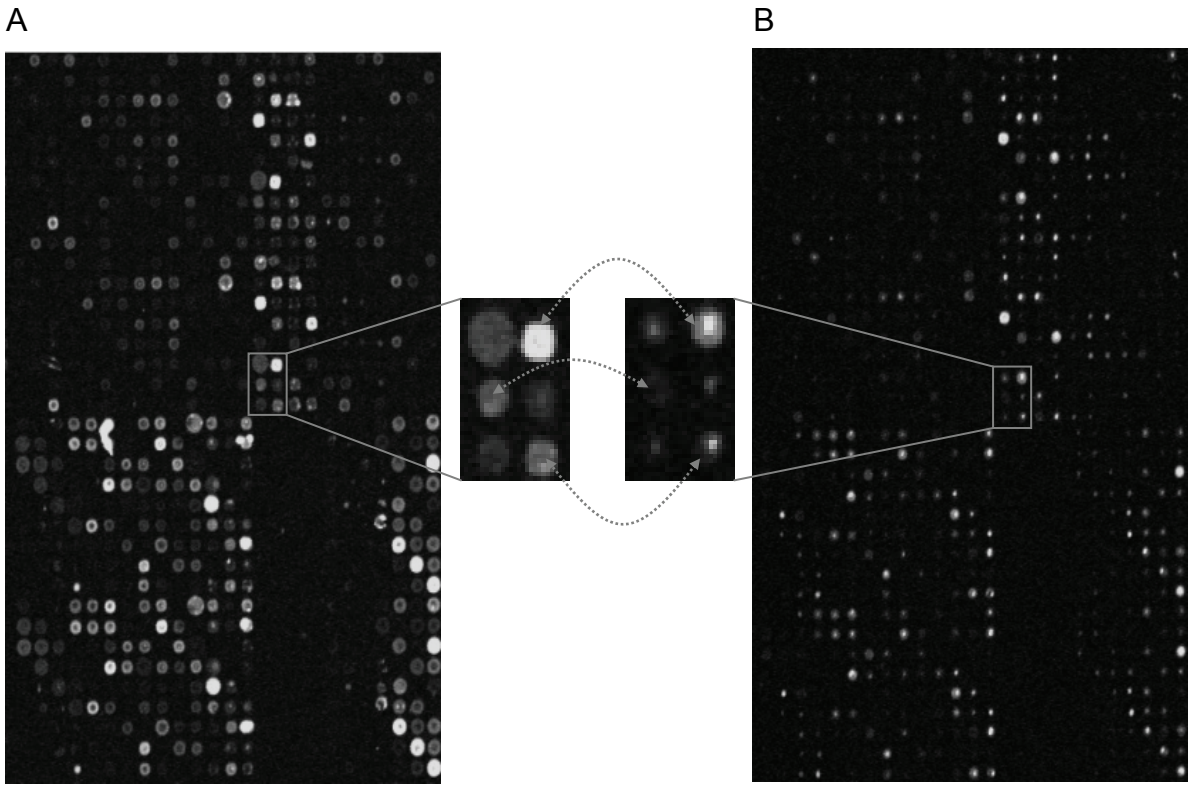


Fig. 2. Hybridisations to *B. subtilis* microarrays made from unpurified PCR products (cy5 signal of the *ccpA* mutant at the transition stage of growth, unpublished data). Decreasing slide quality between the first (panel A) and the third spotting run (panel B) is depicted on the same fragment of the DNA chip. The same part of both slides has been zoomed-in to highlight the differences in spot morphology (upper and lower arrows) and signal intensities (middle arrow).

There are many possible ways to explain the occurrence of this phenomenon:

(i) the spotting buffer used for the *B. subtilis* chips contains DMSO and it is very well possible that DNA dissolved in DMSO-based buffers may aggregate if the DMSO concentration exceeds a certain threshold, which may especially occur upon freeze-thaw cycles of the source PCR plates between the spotting events;

(ii) DNA binding to the plastic surface of the plate after prolonged storage could take place;

(iii) aminated DNA (amino linker present in the first round PCR) may account for cross-reaction with the compounds of the spotting buffer;

(iv) drying-dissolving cycles between spotting events could increase the concentration of contaminants (peptides, amino acids, salts) that in turn might lead to an inefficient immobilization of the DNA due to the competition of functional groups on the surface of the glass.

Interestingly, only part of the material showed the above mentioned behaviour that may be related to the actual concentration of individual amplicons, as the low concentrated ones have a tendency to be more prone to form small dots instead of regular spots. Still, attempts have been undertaken to improve the quality by using slides with different chemistry of coating, changing spotting conditions and testing various processing methods after spotting, all of which failed to provide acceptable hybridisation signals. All in all, the results obtained with the microarrays containing the spots of non-optimal morphology would give little confidence that the measured signal intensities are truly due to specific hybridisation events on the chip. Therefore, no further experiments were performed with the PCR-based microarrays owing to their inferior quality and a new effort was initiated to amplify the whole set of amplicons and this time the purification step was included. In our hands, the yields of the purified PCR products were not sufficient to obtain saturated DNA concentrations on the microarrays. Especially in the case of smaller amplicons (< 300bp), a considerable loss of material after purification was experienced.

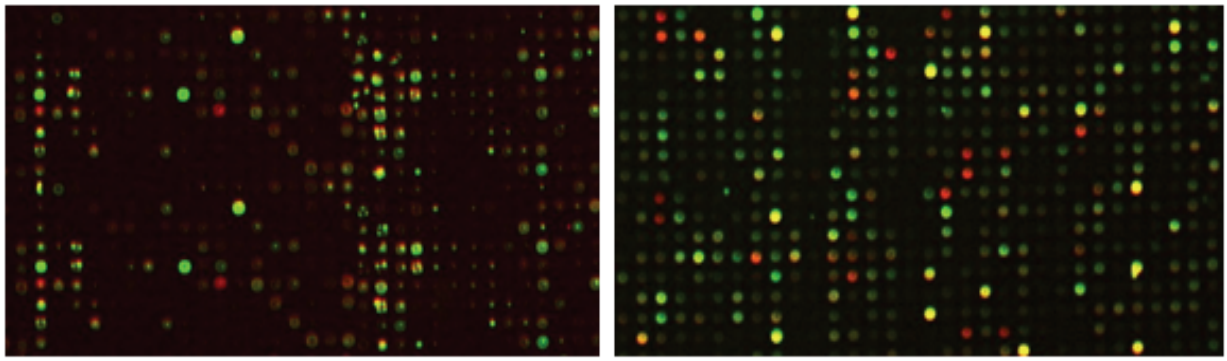


Fig. 3. Comparison of hybridization profiles to *B. subtilis* DNA microarrays prepared from PCR products (left) and oligonucleotides (right).

At this stage, it was decided to design and purchase the whole set of ready-to-spot oligonucleotides (~70 bp) since the market price of oligonucleotides decreased considerably at that time (end of 2004). A part of oligonucleotide material was tested in different spotting buffers and on slides with diverse coating for selection of the most optimal conditions. As shown in Fig. 3 the oligonucleotide microarrays gave a much better hybridization outcome as compared to the amplicon chips. The former ones were finally used for the CcpA (chapter 5 and chapter 6) and the lactate stress (chapter 4) transcriptome experiments. Only in the case of the secretion stress (chapter 3) whole genome analysis was performed with commercially available nylon membrane macroarrays (Sigma Genosis), as this experiment was performed before the microarrays for *B. subtilis* became fully operational in our department.

Principles of microarray experiments

Experimental procedures of a microarray experiment

Typical steps of a DNA microarray experiment include:

(i) cultivation of cells and RNA preparation from the reference and sample cells (for example of a wild-type and knock-out strain or strains grown under normal and stress conditions);

(ii) reverse transcriptase reaction to synthesize cDNA from the extracted RNA;

(iii) labeling of the cDNA;

(iv) hybridization of labeled cDNA to target DNA on a slide (usually each gene is spotted *in duplo*);

(v) collection of the hybridization signals corresponding to both samples by scanning the slide;

(vi) quantification of the signals (conversion of images into numerical information) and data collection;

(vii) data manipulation and biological interpretation of the results (Fig. 4).

Although the whole procedure seems to be quite straightforward, in the early 2000's, there was still little experience available about the experimental pitfalls of this technology and that is why each and every single step required fine-tuning in order to optimize the final outcome. Similarly to the process of the slide fabrication presented above, there were many pitfalls along the way to obtain hybridization signals of sufficient quality for further downstream processing. Since there are many critical steps at all stages in the microarray experiment it is sensible to incorporate check-up points to ensure that samples loaded onto expensive arrays do contain the adequate amount of correctly labeled material. Already at the very beginning of the procedure one has to stay alert as the time of culture sampling should be as short as possible so that the isolated RNA would be representative of the population present, for example at a specific growth phase. To this end, samples for RNA isolation should be snap-frozen immediately after collection to stop transcription and prevent degradation of unstable RNA moieties. In addition, it has to be noted that most of the bacterial mRNAs have a half-life in minutes and that is why solutions such as RNAlater (Ambion, Qiagen), which stabilize the mRNA population may be considered, especially when an immediate isolation of RNA is not feasible. It is recommended to measure the concentration as well as the integrity of the isolated RNA by using spectrophotometry and gel electrophoresis, respectively. In the former case the NanoDrop (Thermo Scientific) is a very useful machine as it can measure as little as 0.5 μl sample, thus opposite to classical spectrophotometers there is no need to dilute samples. Since running RNA agarose gels is quite an elaborate procedure, the total RNA sample can be also easily assessed by controlling the integrity of 16S and 23S units of bacterial ribosomal RNA in the Agilent 2100 Bioanalyzer, which is a very rapid and convenient method. Both the NanoDrop and the Agilent machines were not only routinely used for

evaluation of RNA quality in the experiments presented in this thesis, but they were also employed for comparison of different commercially available RNA isolation kits and RNA isolation methods.

The following step in the microarray procedure is the conversion of the RNA to cDNA by using the reverse transcriptase reaction during which fluorescent dyes (such as most commonly used Cy3 and Cy5) or, in the case of nylon DNA macroarrays a radioactive label, are incorporated. This so-called direct labeling method was tried in the testing phase of *B. subtilis* microarray development, but it turned out not to be very successful due to the unequal incorporation efficiencies of Cy3- and Cy5-labeled nucleotides (dUTPs). Notably, the Cy5 dye is bulkier than Cy3 due to which it is more difficult to incorporate it into cDNA during the reverse transcriptase reaction. In addition, the Cy5 dye is more sensitive to light (photo-bleaching) and ozone, which under improper handling during labeling, hybridization and scanning, may lead to a considerable loss of red signal during the whole procedure. The problem of Cy5 incorporation can be circumvented by an indirect, two-step labeling method. In the first step aminoallyl-dUTP is integrated in the cDNA and thereafter Cy-dyes are coupled to this aminoallyl-group resulting in equal and higher efficiencies of the dye incorporation when compared to the direct labeling. At this stage the quality check-up should be implemented to evaluate the concentration of cDNA and the dye-incorporation rates. Also in this case the NanoDrop spectrophotometer is of help since it can very precisely detect single stranded DNA concentrations and the fluorescent dye-labeling density of any nucleic acid samples. In case sufficient quantities of labelled cDNA are obtained (usually 50-100 pmol of incorporated dye), equal amounts of Cy3 and Cy5 probes are combined and hybridized to slides. It is common practice to perform a dye-swap experiment (i.e. to label the same probe once with Cy3 and once with Cy5) to compensate for labeling differences between the two dyes. Hybridization begins when the solution containing the targets is placed onto the slide and thereafter covered with a cover slip and placed in a humidified chamber or in a special hybridization cassette in case an automatic hybridization station is used. Hybridization protocols may vary a lot and depending on the type of the slide coating and the GC-content of the organism's genome, specific buffers, different incubation temperatures and times as well as washing conditions are practised. After hybridization, slides are washed to reduce the effects of nonspecific hybridization signals and scanned to quantify the amount of target bound to the DNA probes. The underlying principle of this procedure is that a laser excites the fluorescent dyes causing the release of photons, the number of which is compared between the Cy3 and Cy5 spectra. It is important to use appropriate scanning settings to assure that there is no saturation present and that the linear and similar range of the scanner is used for both dyes. Scanning generates gray-scale images which are later processed in especially designed software (for example ArrayPro 4.5 from Media Cybernetics inc.) to quantify signal intensities of spots and estimate the background from both, Cy3 and Cy5, channels. The scanned images of microarrays are often visualized in the form of artificial colour spots on a black background. Green and red spots indicate that a particular gene was

more abundant in the sample labeled with Cy3 and Cy5, respectively. Yellow spots mean that there is equal level of transcription between the two populations of the test and the reference strains and black spots represent genes with no activity under the tested conditions.

Data analysis and validation

Once the DNA microarray technology is well established and implemented, the subsequent data interpretation and data mining require a major and most important effort. This is caused by the fact that microarray experiments generate massive amounts of data. To give an example, in the time series transcriptome analysis described in chapter 5 of this thesis, in which the CcpA mutant strain was compared to its wild-type counterpart at four individual growth phases with four biological replicates per time point and two technical replicates (as each gene is represented twice on the slide), more than a quarter of a million measurements were collected. There are many extensive reviews available that discuss in detail a variety of sophisticated approaches related to microarray data analysis (see for example [143,203]). For this reason only a brief outline of this topic will be given here. In principle, several steps can be distinguished in the process.

(i) The analysis begins with the conversion of raw scanner images (usually 16-bit TIFF files) into numbers representing gene expression levels followed by data pre-processing and normalization.

(ii) During pre-processing the background signal is corrected to subtract the local background from the actual spot intensity. Furthermore spots with poor-quality and low-intensity are filtered out. Low-intensity spots, which are much more variable, should be removed to prevent artefacts.

(iii) The aim of normalization is to minimize variations originating from non-biological influences (systematic errors) such as unequal quantities of starting RNA, differences in labeling and signal detection during scanning. Non-linear normalization methods, such as LOWESS (locally weighted scatter plot smoothing), are considered to be superior to global or non-global methods that estimate the intensity of each spot/gene as a percentage of the total signal on the microarray or use a certain set of uniformly expressed genes (special controls or house-keeping genes), respectively. The LOWESS method takes gene intensity and spatial distribution into account that (may) occur when individual regions of the array are printed by different pins resulting in spatial patterns with lower of higher hybridization signals [204].

(iv) Thereafter, the ratios (usually log-transformed) of the average normalized intensities are generated and scrutinized by statistical methods in order to identify genes with significantly altered expression levels. Since in most cases microarray experiments are performed in replicates, a confidence level can be assigned to the gene expression ratios. A commonly used method is the determination of p-values in the t-test. Different variants of the t-test can be found in many statistical

software packages. An example of such software is Cyber-T that employs statistical analyses based on simple t-tests that use the observed variance of replicate gene measurements across replicate experiments, or regularized t-tests that use a Bayesian estimate of the variance among gene measurements within an experiment [12]. This software was applied in the analyses presented in all the experimental chapters in this thesis. As to the simple t-test approach, it has to be taken into account that large numbers of genes are simultaneously tested in a microarray experiment (so called multiple testing). By applying a standard confidence level of 5% (p=0.05) a false positive rate of 5% has to be accepted, meaning that for a 4000-gene array, 200 genes would be mistakenly found as significantly expressed. Therefore, corrections methods are needed to deal with multiple testing. One of the best known is the Bonferroni correction [209] in which each value is multiplied by the number of tests (thus in the example given above the p-value of 0.05 should be multiplied by 4000). An alternative approach is the calculation of the False Discovery Rate (FDR) in which the q-value represents the level of significance. The q-value belonging to a measurement describes the proportion of false positive measurements relative to the total number of measurements [84].

As part of the statistical analysis, a clustering approach can be used to group the genes with similar expression patterns [161]. By such an approach the function of unknown genes may be discovered, if they are co-regulated with known genes in response to a specific stimulus. Several clustering algorithms including hierarchical clustering, self-organizing maps, k-means or principal component analysis can be used, depending on the type of expression data [203]. A wide variety of different commercial (for example GeneSpring or Spotfire) as well as open-source (for example CLUSTER, TREVIEW and Genesis; the latter was widely used in the transcriptome analyses presented in this thesis) software packages are available to perform and visualize the outcome of clustering procedures.

(v) To enable the comparison and verification of results, it is crucial to share data originating from various microarray experiments. Therefore, it is a common requirement of scientific journals to submit the results to the databases such as the Gene Expression Omnibus (GEO) [13] or ArrayExpress [24], which archive and distribute data from all microarray platforms. Both databases use the annotation standard of Minimum Information About a Microarray Experiment (MIAME) that describes the minimum information required to ensure that microarray data can be easily interpreted and that results derived from its analysis can be independently verified [23].

(vi) Furthermore, an extensive data mining process can be started with the aim of defining functional categories and pathways affected under the conditions investigated. Also here a large number of tools has been developed [123]. For the

analyses reported further in this thesis, the FIVA software (Functional Information Viewer and Analyzer) has been employed to characterize the affected biological processes by the identification of overrepresented functional categories in clusters of differentially expressed genes [20].

(vii) Finally the results obtained can be submitted to, and compared with, the data available in general and specific databases, such as PRODORIC [81] or DBTBS [226].

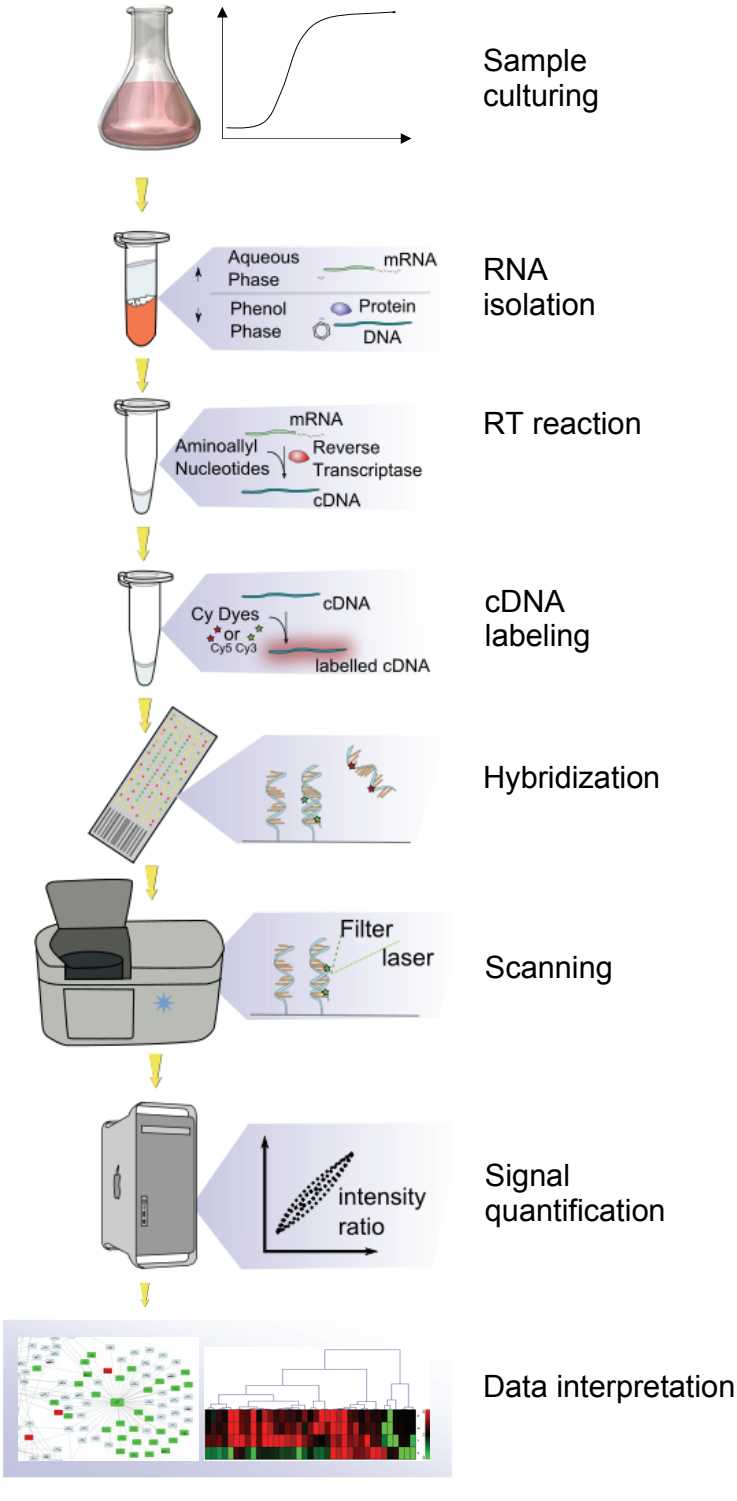


Fig. 4. Steps of a microarray experiment (adapted from wikipedia.org). See text for details. For reasons of simplicity the processing of only one of the two samples (for example wt or mutant) is depicted here.

As outlined above, a DNA microarray experiment involves a multistep assay due to which it is prone to unintentional errors that may happen during fabrication of slides, the array experiment itself and data analysis. To minimize the chance of unreliable results, array experiments need to be validated as well as possible. This can be achieved either by comparing the transcriptome data to the information available in literature and public expression databases or by laboratory-based approaches [40]. In the latter case, methods such as real-time reverse transcriptase PCR, Northern blot, primer extension, reporter gene systems (β-gal or GFP for example) and electrophoresis mobility shift assays may be implemented to collect supporting evidence for the array results. Taking into consideration that a microarray experiment generates expression profiles of thousands of genes and that classical "old-fashioned" experiments are usually laborious, it is difficult to decide how many genes should be actually selected to offer a reasonable level of confidence of the transcriptome data.
In this respect, if there is enough *in-silico* evidence and reliable statistics to support the validity of the data, one may argue to omit confirmatory studies [210].

Transcriptome studies of *B. subtilis*

In the recent decade various reviews have been published in relation to applications of microarray technology in bacterial systems [54,161,274]. In this chapter a special focus will be paid to the use of gene expression profiling of *B. subtilis*. The first reports on genome-wide transcriptional profiling of *B. subtilis* become available in 2000 [67,273] and up till now over 100 publications on this topic can be found in the PubMed database. During preparation of this introduction a comprehensive review on microarray studies in *B. subtilis* became available in which 105 publications (published till April 2009) were summarized and divided into seven themes:

(i) effects of gene deletion or overexpression,
(ii) effects of overexpression of heterologous proteins,
(iii) comparison of gene expression under aerobic vs. anaerobic conditions,
(iv) effects of temperature changes,
(v) effects of transported molecules,
(vi) effects of limitations and stress conditions, and
(vii) other transcriptome studies [131].

This categorization could be made more general as factors such as temperature change, oxygen limitation as well as overexpression of proteins and exposure to certain molecules can be also listed under the category "stress conditions". That is why Tables 1 and Table 2 show a selection of publications categorized in 2 broad classes: transcriptional regulation and response to environmental changes/stress, a division that allows defining regulons and stimulons, respectively. The former category will also include experiments that describe

expression patterns that are specific to growth or developmental stage. A regulon is a set of expression units sharing a common regulatory binding motif for a specific TF, whereas a stimulon describes a group of differentially expressed genes in response to a particular environmental change. A typical approach to identify regulons involves a comparison of the strain carrying an inactive form (knock-out mutant) of a regulatory protein or an overproducing strain against its wild-type counterpart. A detailed elaboration on all the transcriptome analyses as presented in Tables 1 and 1 falls outside the scope of this thesis, due to which only some general aspects will be pinpointed below.

Table 1. Regulons of *B. subtilis* studied with microarrays.

TF	Function	ref*
	Global regulators	
AbrB	transcriptional pleiotropic regulator of transition state genes	[87]
CcpA	transcriptional regulator mediating carbon catabolite repression and nitrogen metabolism	[18,150,154, 181,276]
CodY	repressor of early stationary-phase genes	[179]
ComK	competence transcription factor ComK	[16,86,188]
Spo0A	master regulator for entry into sporulation	[178]
TnrA	global regulator that responds to the availability of nitrogen sources	[245,278]
Spx	global transcriptional regulator of the oxidative stress response	[38]
	Sigma factors	
SigB	general stress response	[193,199]
SigD	flagellar synthesis, motility, chemotaxis, autolysis	[224]
SigE	sporulation mother cell-specific (early)	[68]
SigF	sporulation forespore-specific (early)	[67]
SigH	transition from exponential to stationary growth	[26]
SigM	cell envelope stress, essential under high concentrations of salt (ECF)	[58]
SigV	regulatory overlap with SigX, SigY and SigZ (ECF)	[284]
SigW	detoxification, production of antimicrobial compounds (ECF)	[32]
SigY	nitrogen starvation	[33]
ECF's	seven extracytoplasmic function (ECF) sigma factors	[8]
SigEFGK	compartment-specific sporulation sigma factors	[233]
	two-component systems (TCS) and other regulators	
Zur	regulation of genes involved in zinc uptake	[76]
CcpN	repression of gluconeogenic genes	[239]
CysK	global negative regulator of genes involved in sulfur metabolism	[3]
SinR	regulation of biofilm formation	[39]
Fnr	adaptation to low oxygen tension	[206]
Fur	response to iron starvation	[11]
LmrA	multidrug resistance	[277]
RelA	GTP pyrophosphokinase (stringent response)	[64]
ScoC/hpr	transcriptional repressor of sporulation and extracellular proteases genes	[30]
TCS	24 two-component regulatory systems	[129]
ComXPA	ComX-ComP-ComA quorum-sensing pathway	[41]
DegUS	TCS involved in degradative enzyme and competence regulation	[159]

*The PubMed database was searched with the different combinations of the following primary keywords: "(micro)array(s)", "macroarray(s)" "transcriptome/transcriptomic analysis/profiling", "genome-wide analysis" and "subtilis"

Overall, more than 60 TFs, including global regulators, sigma factors, regulators of TCSs, have been investigated (Table 1) and their influence on the transcriptome ranged from less than 5 genes (LmrA, CcpN, SigY) to over 400 genes (SpoA, ComK, CcpA, ScoC). In many cases the microarray approach was supplemented with additional data obtained by techniques such as chromatin immunoprecipitation (so called ChIP-on-Chip), proteomics or bioinformatics. The chromatin immunoprecipitation allows identification of DNA regions to which a TF binds *in vivo* and this approach in combination with transcriptional profiling, gel electrophoretic mobility shift assays, and bioinformatics has proven to be a powerful method that enabled researchers to assign 103 genes to the Spo0A regulon in addition to 18 previously known members [178]. Combined proteome and transcriptome analysis offers good opportunities to correlate gene expression levels to the presence of corresponding proteins in the cells, but it has to be noted that even though the results of both methods can be in good agreement, usually less targets are found by proteomics. For example, a transcriptomic analysis detected 66 glucose-repressive genes of which only 17% were confirmed by 2D gel electrophoresis [276]. Bioinformatics is commonly employed, since it cannot only facilitate the analysis (concerning issues such as data mining, visualization and statistics) of huge amounts of data originating from microarray analyses, but also helps in deduction of new or additional consensus sequence(s) recognized by a particular TF. It has to be noted that the presence of an upstream regulatory sequence has to be cautiously interpreted. *B. subtilis* contains 1062 putative ComK-binding sites (K-boxes) in its genome but based on the microarray experiments it was found that only approximately 8% of the genes containing a K-box in the promoter region are really regulated by ComK [86]. Oppositely, in case of CcpA a large group of genes without any clear binding motive were regulated by this global regulator, which only partially can be attributed to indirect effects [154]. Similarly, it was also shown that some genes associated with Fur boxes were not regulated by Fur, whereas other genes which did not contain a conventional Fur consensus were indeed regulated [11]. As presented in Table 2 *B. subtilis* was "tortured" under a wide variety of stress conditions. Some of these conditions, such as copper limitation, triggered differential expression of only a few genes [35]. On the other hand, the exposure of cells to erythromycin or heat caused a massive reprogramming of the expression network involving over 1200 and 400 genes, respectively [97,146]. With some exceptions, many different stress signals led to an induction of the general stress genes, subject to the control of SigB. These genes code for proteins that perform different protective functions including non-specific protection against oxidative stress, maintenance of redox balance under energy depletion, adaptation to temperature, acidic changes, alkaline shifts and osmotic stress, reorganization of cell-wall/membrane components and secretion of toxic compounds. The importance of the general stress response system is stressed by the fact that normally growing cells devote only 1% of their translation capacity to this system, whereas under stress up to 40% of the capacity is engaged to synthesize stress proteins [93]. In addition to general stress proteins

also stress-specific systems, which involve the expression of a small subset of genes, play a crucial role in defense against unfavorable conditions.

Table 2. Stimulons of *B. subtilis* studied with microarrays.

Type of stress	Effect/influenced regulons and/or genes	ref*
Copper limitation	*ycnJ* upregulation	[35]
Calcium excess	Influence on sporulation, flagella and biofilm formation	[189]
Iron starvation	Genes linked to stringent-response regulation and amino acid biosynthesis associated with pathways essential for bacillibactin production	[11,168]
Manganese excess	Perturbation of cellular iron pools (increased Fur activity). Activation of SigB and TnrA regulons	[83]
Nitrogen starvation	ComG, ComF, ComE, nin-nucA, RecA and ComK regulons	[113]
Phosphate starvation	New members of the PhoP regulon: *yfkN*, *yurI*, *yjdB*. Activation of SigB regulon	[5]
Glucose starvation	Extensive reprogramming of gene expression: (i) SigB regulon, (ii) carbon metabolism and utilization	[130]
Glucose repression	66 glucose-repressive genes, 22 of which were at least partially under CcpA-independent control	[276]
EtOH, heat, salt	Confirmation of known general stress genes and detection of 63 additional genes controlled by SigB	[193]
Anaerobic growth	Several hundred genes were induced or repressed involved in carbon metabolism, electron transport, iron uptake, antibiotic production and stress response	[273]
TrxA depletion	Thiol-disulfide oxidoreductases (maintaining the thiol redox state of cells) induction of oxidative stress response and sulfur utilization. Influence on sporulation and competence-related genes	[182,227]
Salt stress (NaCl)	1.2 M NaCl: (i) induction of SigW regulon (ii) DegS/DegU, (iii) a high-salinity-mediated iron limitation, and (iv) a repression of chemotaxis and motility genes	[232]
Sulfate or methionine	Repression of S-box regulon in the presence of methionine. Induction of genes encoding transporters *yhcL*, *ytmJKLMN* and *yxeMO*	[9]
Aminoacid availability	Repression of genes involved in amino acid biosynthesis, sporulation, and competence, in casamino acids containing medium	[160]
Heat (48°C)	Activation of SigB, HrcA and CtsR regulons	[97]
Cold shock (15°C and 18°C)	18°C: Repression: biosynthesis of amino acids, nucleotides and coenzymes. Strong induction of fatty acid desaturase (*des*) and cold-shock genes *cspB*, *cspC* and *cspDC* and enzymes involved in degradation of branched chain amino acids. 15°C: Induction of the SigB regulon, SigF, SigE and SigG and RapA/PhrA and Opp. Reductions in major catabolic (glycolysis, oxidative phosphoryl. ATP synthesis) and anabolic routes (biosynthesis of purines, pyrimidines, haem and fatty acids). Repression of part of the SigW regulon and genes involved in chemotaxis and motility	[14,28,119]
Alkaline stress	Induction of SigW regulon	[269]
Acid / alkaline (pH 6 & pH 9)	pH 6 activation of acetoin production (*alsDS*), dehydrogenases (*adhA, ald, fdhD,* and *gabD*), and decarboxylases (*psd* and *speA*), malate metabolism (*maeN*), metal export (*czcDO* and *cadA*), oxidative stress and SigX regulon. pH 9: activation of arginine catabolism (*roc*), glutamate synthase (*gltAB*), K(+)/H(+) antiporter (*yhaTU*), cytochrome oxidored. (*cyd*, *ctaACE*, and *qcrC*), SigH, SigL, and SigW regulons. greater genetic adaptation was seen at pH 9 than at pH 6	[270]
Salicylic acid	Induction of SgB, CtsR and Spx regulons and phenolic acid decarboxylases PadC and BsdBCD (YclBCD)	[57]

Sorbic acid	Strong derepression of the CcpA, CodY, and Fur regulon and induction of TCA cycle genes, SigL- and SigH-mediated genes, fatty acid biosynthesis (*fab*) genes and BkdR-regulated genes and the stringent response	[241]
Type of stress	**Effect/influenced regulons and/or genes**	**ref***
Phenol and catechol	Activation of the HrcA and CtsR heat-shock regulons, Spx, PerR and FurR regulons. Derepression of catabolite-controlled genes and strong induction of the *yfiDE* operon	[237]
Thiol oxidant diamide	Activation of PerR (oxidative stress) and CtsR (heat shock) regulons, putative regulators and proteins protecting against toxic elements and heavy metals. Repression of genes of stringent response	[142]
Furanone	Inhibition of growth, swarming, and biofilm formation. Activation of class III heat shock genes, fatty acid biosynthesis, lichenan metabolism	[208]
Chloramphenicol Erythromycin Gentamycin	Activation of genes encoding transport/binding proteins, protein synthesis, and metabolism of carbohydrates and related molecules	[146]
Vancomycin	Induction of SigW regulon	[34]
Daptomycin Friulimicin B	Induction of ECF sigma factors. LiaRS is strongly induced only by daptomycin, indicative of different mechanisms of action of these two antibiotics	[264]
Methylglyoxal and formaldehyde	Induction of Spx, CtsR, CymR, PerR, ArsR, CzrA, CsoR, SigD regulons and SOS response (DNA damage)	[184]
Chromanon 2-MHQ (antimicrobials)	Chromanon- induction of the HrcA, CtsR, and Spx regulons. 2-MHQ - induction of oxidative stress response, dioxygenases and oxidoreductases (*yfiDE*, *ydfNOP*, *yodED*, *ycnDE*, *yodC*, and *ykcA*)	[185]
Protein overproduction/ secretion (AmyQ)	Induction of CssRS regulon. Repression of several genes of the sporulation pathway and induction of a group of motility-specific (*sigD*-dependent)	[109,155]
Insoluble heterologous protein overproduction (PorA)	Induction of heat shock genes of class I (*dnaK*, *groEL* and *grpE*) and class III (*clpP* and *clpC*), ribosomal proteins RpsB and RplJ, pyrimidine and purine synthesis enzymes and ribosomal protein genes	[118]

*The PubMed database was searched with the different combinations of the following primary keywords: "(micro)array(s)", "macroarray(s)" "transcriptome/transcriptomic analysis/profiling", "genome-wide analysis" and "subtilis"

Such specific stress responses were observed by heat shock (governed by HcrA), cold shock (DesRK), protein overproduction/aggregation stress (CtsR, CssRS), oxygen stress (PerR, FNR, ResRD), phosphate starvation (PhoRP) and nitrogen limitation (TnrA, GlnR). Under many stress situations the interplay of both defense systems (specific and general) is crucial to stand up to the encountered adverse conditions.

Concluding remarks and perspectives

The explosion in available genome sequence data of the last decade resulted in the subsequent development of DNA microarray technologies that allowed analysis of whole genome gene expression in a single experiment. DNA microarrays have become a common tool in many areas of bacterial research with special attention paid to responses to stress conditions and identification of entire regulons. Their application has extended to a broad

range of research fields such as microbial pathogenesis, epidemiology, ecology, pharmacology. Last but not least, microarrays gain more and more popularity in diagnostics. It is worth mentioning that different transcriptomics technologies have been used in the whole-genome analyses of *B. subtilis* and other bacteria, including nylon macroarrays, in-house spotted glass microarrays containing PCR products or oligomers and also commercially available Affymetrix chips. However, in the near future it can be foreseen that these "standard" expression microarrays will be gradually replaced by the new emerging technologies. There is an ongoing improvement in the microarray platforms, an example of which is the development of tiling DNA microarrays. This kind of microarrays are build-up with a set of overlapping oligonucleotide probes that span the entire genome at high resolution [229,259]. The first results obtained with tiling arrays have been reported recently and show that this technology has high potential as compared to traditional methods to gain even better insights toward a comprehensive understanding of the transcriptome of *B. subtilis* [205]. A very recent addition to the array of possibilities for transcriptome analysis is the use of ultra-deep (ultra-high-throughput) sequencing [71]. The first publications of this powerful technology have emerged and undoubtedly this technology holds great promise for even more reliable and faster high throughput analysis of gene expression in cells.

Chapter 3

Production and secretion stress caused by overexpression of heterologous α-amylase leads to inhibition of sporulation and a prolonged motile phase in *Bacillus subtilis*

Andrzej Lulko, Jan-Willem Veening, Girbe Buist, Wiep Klaas Smits, Evert Jan Blom, Aaron Beekman, Sierd Bron and Oscar Kuipers

ABSTRACT

Transcriptome analysis was used to investigate the global stress response of the Gram-positive bacterium *Bacillus subtilis* caused by overproduction of the well-secreted AmyQ α-amylase from *Bacillus amyloliquefaciens*. Analyses of the control and overproducing strains were carried out at the end of exponential growth and in stationary phase, when protein secretion from *B. subtilis* is optimal. Among the genes that showed increased expression were *htrA* and *htrB*, which are part of the CssRS regulon that responds to high-level protein secretion and heat stress. The analysis of the transcriptome profiles of a *cssS* mutant compared to the wild-type, under identical secretion stress conditions, revealed several genes with altered transcription in a CssRS-dependent manner, for example *citM*, *ylxF*, *yloA, ykoJ* and several genes of the *flgB* operon. However, a high affinity CssR-binding was only observed for *htrA* and *htrB*, and possibly for *citM.* In addition, the DNA macroarray approach revealed that several genes of the sporulation pathway are downregulated by AmyQ overexpression, and a group of motility-specific (σ^D-dependent) transcripts were clearly upregulated. Subsequent flow cytometric analyses demonstrate that upon overproduction of AmyQ as well as of a non-secretable variant of the α-amylase, the process of sporulation is severely inhibited. Similar experiments were performed to investigate the expression levels of the *hag* promoter, a well-established reporter for σ^D-dependent gene expression. This approach confirmed the observations based on our DNA macroarray analyses and led us to conclude that expression levels of several genes involved in motility are maintained at high levels under all conditions of α-amylase overproduction.

Introduction

The Gram-positive bacterium *Bacillus subtilis* is capable of secreting high amounts of endogenous proteins into the extracellular medium [268]. Therefore, this bacterium and its relatives are often exploited as hosts for the production and secretion of heterologous industrially interesting enzymes. Secretion of heterologous proteins in large quantities has been shown to lead to the unfavourable condition of protein misfolding and subsequent degradation [216]. In *B. subtilis*, accumulation of misfolded proteins at the membrane-cell wall interface is sensed by the CssRS two-component system, which consists of the membrane-embedded sensor kinase CssS and the response regulator CssR [108]. This system responds to high-level protein secretion and heat stress by phosphorylation of CssR which, in turn, activates transcription of the monocistronic *htrA* and *htrB* genes [48]. The CssRS system has also been shown to regulate the expression of its own operon [48,129]. HtrA and HtrB are membrane-bound serine proteases whose major functions are degradation of misfolded and aggregated proteins. Expression of both proteases is induced upon excessive expression of secretory proteins or a temperature increase [187]. Thus, the CssRS regulon forms a quality control and defence system when cells are confronted with secretion stress. The members of the CssRS quality control system become involved at that stage and take action in two possible manners, either by acting as a chaperone aiding refolding or preventing unfolding, or by HtrA/B-dependent proteolysis.

Recently, proteomic studies have been performed to investigate the global effects of secretion stress [7]. These studies showed that HtrA is also present in the culture supernatant and, in addition to its protease activity, can act as a molecular chaperone. In addition, global effects of secretion stress in *B. subtilis* were also investigated by transcriptome analysis during the exponential growth phase. This study, however, only revealed differential expression of relatively few genes, including the genes encoding the HtrA and HtrB proteases [109].

Because most of protein secretion in *B. subtilis* takes place during the onset of the stationary phase of growth [21,99], we set out to examine the global transcriptional response of *B. subtilis* under secretion stress conditions at this specific growth phase. Secretion stress was applied by overproducing the well-secreted AmyQ α-amylase from *B. amyloliquefaciens*. Besides examining secretion stress in wild-type cells, we compared transcriptome profiles of a *cssS* mutant strain under conditions of high-level AmyQ production. In this work, we have identified and verified putative novel members of the CssRS regulon and we dissected direct and indirect effects of α-amylase overproduction and protein secretion in stationary phase cultures. Our study reveals that upon overproduction of a non-secreted α-amylase, as well as the secreted wild-type variant, the process of sporulation is severely inhibited. In addition, we show that expression levels of genes involved in motility are maintained at high levels under all conditions of α-amylase overproduction.

Materials and methods

Bacterial strains and growth conditions

B. subtilis 168 with an integrated spectinomycin (*sp*) marker in the *pks* locus (*B. subtilis* 168::*sp*) [86] and *B. subtilis* 168 *cssS*::*sp* [108] were used for transcriptional analyses (see Table 1). *B. subtilis* 168::*sp* was used to avoid possible effects on the transcriptional profiling due to the presence of the spectinomycin gene alone, as was shown by Hamoen *et al*., and to be able to compare the obtained data of this strain with that of the *cssS* insertion mutant. Both strains were transformed with either the pUB110 (empty vector) [48] or the pKTH10 plasmid (pUB110 derivative containing the *amyQ* α-amylase gene of *B. amyloliquefaciens)* [192].

Table 1. Bacterial strains and plasmids.

Strains and plasmids	Genotype	Reference
B. subtilis		
168	trpC2	[137]
168::sp	168, *pks*::*sp*, Sp^r	[86]
IIA-gfp	168, P_{spoIIA}-gfp, Cm^r	[256]
hag-gfp	168, P_{hag}-gfp, Cm^r	This work
IIA/E	168, P_{spoIIA}-gfp, Cm^r; pUB110	This work
IIA/Q	168, P_{spoIIA}-gfp, Cm^r; pKTHM10	This work
IIA/QAla	168, P_{spoIIA}-gfp, Cm^r; pKTHM102	This work
hag/E	168, P_{hag}-gfp, Cm^r; pUB110	This work
hag/Q	168, P_{hag}-gfp, Cm^r; pKTHM10	This work
hag/QAla	168, P_{hag}-gfp, Cm^r; pKTHM102	This work
CB100	*sigD*::cat, Cm^r	[164]
Sik243	*spo0A*::Em, Em^r	[111]
BV2001	*cssS*::*sp*, Sp^r	[108]
Plasmids		
pUB110	Km^r	
pKTH10	Km^r , pUB110 derivative containing the α-amylase gene (*amyQ*) of *B. amyloliquefaciens*	[192]
pKTHM10	Km^r , pKTH10 derivative	[283]
pKTHM102	Km^r , pKTHM10 with the Ala-rich signal peptide of AmyQ	[283]

B. subtilis strains were grown in TY medium (1% tryptone, 0.5% yeast extract and 1% NaCl) at 37°C, 250 rpm containing the appropriate antibiotics. Antibiotic concentrations were: kanamycin (10 μg/ml), spectinomycin (100 μg/ml) and chloramphenicol (5 μg/ml).

Strain construction

To construct plasmid pGFP-hag, a PCR with primer pair hag-F and hag-R (see Table 2) was performed, using chromosomal DNA of *B. subtilis* 168 as a template. The amplified fragment containing the promoter region of the *hag* gene was subsequently cleaved with *Hin*dIII and

*Eco*RI, and ligated into the corresponding sites of pSG1151, in that way generating a fusion with the *gfpmut1* gene [144]. *B. subtilis* strain hag-gfp was obtained by Campbell-type integration of plasmid pGFP-hag into the chromosome of *B. subtilis* 168. Transformants were selected on TY agar plates containing chloramphenicol (5 μg/ml), after overnight incubation at 37°C. Correct integration was verified by PCR (data not shown).

Table 2. Primer sequences.

Primer	Sequence (5' > 3')	Target fragment	Restriction site
hag-F	GGGATCAAGTGAAGCTTGGAATTGACG	*hag*	*Hin*dIII
hag-R	CGGAATTCCATTTCCTCTCCTCCTTGAATATGTTGTTAAGGCACGTCC	*hag*	*Eco*RI
QE60-cssR-F	CATGCCATGGCATACACCATTTATCTAGTTGAAGA	*cssR*	*Nco*I
QE60-cssR-R	CGGGATCCTGATGACATCATCCTGTAGCCGAAACCGTA	*cssR*	*Bam*HI
QE30-cssR-F	CGGGATCCTTGTCATACACCATTTATCTAGTTGAAG	*cssR*	*Bam*HI
QE30-cssR-R	CCGAGCTCTTATCATGATGACATCATCCTGTAGCCGAAA	*cssR*	*Sac*I
htrA-EMSA-F	AACGGATCAGCCGATACGTT	P*htrA*	None
htrA-EMSA-R	TCATCACGATAGTTATCCAT	P*htrA*	None
htrB-EMSA-F	CGTCAGCAGTTCATTCAG	P*htrB*	None
htrB-EMSA-R	CCATCACGTCGATAATCC	P*htrB*	None
citM-EMSA-F	AGGTCACCTCCTCACCTGAA	P*citM*	None
citM-EMSA-R	CATGAGAAAGCCTAAGATTGCTAAC	P*citM*	None
flgB-EMSA-F	TGTATCGTTCCAGAAAATAAGC	P*flgB*	None
flgB-EMSA-R	TCAGGGCGAGAAATGTAGTTC	P*flgB*	None
ykoJ-EMSA-F	TGCCGATCAAATCAGCAG	P*ykoJ*	None
ykoJ-EMSA-R	TTGTGAGCCCCTCCTTTGT	P*ykoJ*	None
yloA-EMSA-F	CCAGAAGCACAAGCACCATA	P*yloA*	None
yloA-EMSA-R	GTATGTAAACATGCCATCAAACG	P*yloA*	None
pK-F	AATCTATCGACATATCCTGCAA	P*comK*	None
K-FP-R	GGAATTCTTGCGCCGTTCACTTCATAC	P*comK*	None

RNA isolation, preparation of labelled cDNA and hybridization

Cells were grown overnight in 10 ml of TY medium with kanamycin (10 μg/ml) and then diluted to an optical density at 600 nm of 0.1 in 40 ml of TY medium containing kanamycin. Samples for transcriptome analyses were collected at the late exponential growth stage (one hour before the transition point) and 3 hours upon entry in the stationary growth phase. Three independent cultures of each strain were used and cells were sampled for macroarray experiments. RNA was isolated by spinning down cells from 4 ml of culture, subsequent cell disruption, phenol-chloroform extraction, followed by RNA purification with a High Pure RNA isolation kit from Roche, as described previously [86]. RNA was eluted with 50 μl of elution buffer and subsequently quantified with GeneQuant (Amersham) and the RNA integrity was checked on agarose gels. 4 μg of RNA as a template and 1 pmol of ORF specific primers (Eurogentec) were used for the reverse transcriptase reaction with SuperscriptII (Gibco BRL). The detailed protocols of reverse transcription and purification of ^{33}P-labelled cDNA are outlined by Hamoen *et al*. [86]. Labeled cDNA was hybridized to *B. subtilis* Panorama™ Arrays (Sigma-Genosis), according to the manufacturer's instructions. The

membranes were exposed to Cyclone phosphoimager screens (Packard) for approximately 2.5 days and signals were quantified with Array-Pro Analyzer 4.5 (Media Cybernetics).

Data analysis

Duplicate spots were averaged in Array-Pro software (Media Cybernetics, Inc.) and the signal was normalized after background subtraction by calculation of the percentage of total signal per gene using Microsoft Excel. Outstandingly high signals (for example whole genomic DNA control spots which act as a positive control and always show very strong signals) were excluded from calculations of the total signal. Statistical significance of the obtained gene expression ratios was assessed by the Cyber-T program [149]. Genes with a Cyber-T *p*-value lower than 0.05 are discussed in this paper. The array data have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE8014.

SDS-polyacrylamide gel electrophoresis and Western blotting

To detect the precursor and mature forms of AmyQ in cells and growth medium, 1 ml of culture was centrifuged and the cell pellet was washed once with fresh medium. After separation by SDS-PAGE, proteins were transferred to Immobilon-PVDF membranes. The visualization of AmyQ protein was carried out with specific anti-AmyQ antibodies and horseradish peroxidase-anti-rabbit IgG conjugates and ECL immunobloting detection reagents (Amersham).

Amylase activity assay

Samples from culture supernatants were diluted 100 times and were subjected to amylase activity quantification with the EnzCheck Ultra Amylase Assay Kit (Molecular Probes), according to the manufacturer's instructions. Assays were carried out in 96-well microtiter plates in 100 µl total volume using 50 µl diluted supernatants. Reactions contained 200 µg/ml DQ starch substrate and were carried out in 100 mM MOPS at pH 6.9. The degradation of the substrate by amylase yields highly fluorescent fragments and fluorescence was monitored at room temperature every 5 min for 1h with a microplate reader TECAN (GENios) using standard fluorescein filters. Samples were taken from three independent experimental replicates.

Flow cytometry

Cells were 100 times diluted in 0.2 µM of filtered minimal medium [6] and directly measured on a Coulter Epics XL-MCL flow cytometer (Beckman Coulter, Mijdrecht, NL), operating an argon laser (488 nm) as essentially described by Veening *et al* 2005 [256]. For each sample, at least 20,000 cells were analyzed. Data containing the green fluorescent signals were collected by a FITC filter and the photomultiplier voltage was set between 700 and 800 V. Data was

captured using System II software (Beckman Coulter) and further analyzed using WinMDI 2.8 software (http://facs.scripps.edu/software.html). Figures were prepared for publication using WinMDI 2.8 and Corel Graphics Suite 11. To distinguish background fluorescence from GFP-specific fluorescence, the parental strain *B. subtilis* 168 was also analyzed with each flow cytometric experiment.

Construction and purification of 6xHis-tagged CssR

Plasmids encoding CssR protein with a 6xHis-tag were constructed as follows. For C-terminally tagged protein, the *cssR* open reading frame was amplified by PCR using primers QE60-cssR-F and QE60-cssR-R (see Table 2) and Phusion polymerase (Finnzymes). Chromosomal DNA of *B. subtilis* strain 168 was used as a template. The product of this reaction was digested with *Nco*I and *Bam*HI, and cloned into similarly digested pQE60 plasmid (Qiagen), yielding plasmid pQE60-cssR.

To obtain an N-terminally tagged CssR protein, the *cssR* open reading frame was amplified by PCR using primers QE30-cssR-F and QE30-cssR-R. The product of this reaction was digested with *Bam*HI and *Sac*I, and cloned into similarly digested pQE60 plasmid (Qiagen), yielding plasmid pQE30-cssR.

To induce expression, *E.coli* JM109 carrying pQE60-cssR or pQE30-cssR was diluted 1:100 to fresh medium from an overnight culture. At an OD600 of 0.6, expression was induced by the addition of 1mM of IPTG. Cells were harvested after an additional two hours of growth, after which protein was purified as described elsewhere (Smits *et al*, in press). The purity of the protein was estimated to be >95% pure on the basis of Coomassie stained SDS-PAGE gels. After purification, the protein was dialysed against dialysis buffer (20 mM Tris-HCl pH8, 1mM EDTA, 10 mM $MgCl_2$, 0.2 M NaCl, 1 mM beta-mercaptoethanol, 5 mM imidazole, 7% glycerol). The concentration of protein was determined on the basis of A280 using the Nanodrop machine, with 0.1% (1g/L) giving a value of 1.272 (ExPASy Protparam Tool; http://www.expasy.org/tools/protparam.html).

Electrophoretic mobility shift assays

To establish direct binding of CssR to putative target promoters, gel shift experiments were carried out as described before [4]. In our experiments, we observed no significant difference in affinity between C- and N-terminally His-tagged proteins. Subsequent experiments were therefore carried out using only N-terminally His-tagged protein. The fragments obtained corresponded to basepairs -339 to +19 (*htrA*), -335 to +19 (*htrB*), -333 to +26 (*citM*), -191 to +25 (*flgB*), -320 to +0 (*ykoJ*) and -312 to +32 (*yloA*) compared to the start of the open reading frame as annotated in SubtiList R16.1 (http://genolist.pasteur.fr/SubtiList/). As a negative control the promoter of *comK* (-201 to +47) was amplified. All primer sequences are available in Table 2.

Results

Transcriptomics of secretion stress in late exponential and stationary phase cells

AmyQ encodes the α-amylase from *Bacillus amyloliquefaciens* that contains a Sec-type signal sequence. This protein has been shown to trigger a specific secretion stress response when overexpressed in *B. subtilis* [108]. DNA macroarray analysis was used to compare transcriptional profiles of *B. subtilis* 168::*sp* containing plasmid pUB110 (empty vector) with those of *B. subtilis* 168::*sp* containing plasmid pKTH10 (AmyQ overexpression) grown in TY broth. Samples were taken for transcriptome analyses during late exponential (one hour before the transition point) and stationary growth (3 hours after the transition point) phases of growth. Amylase-dependent starch degradation on TY-agar plates and Western blot analysis of the growth media from the cells used for the macroarray experiment, verified the expected AmyQ overproduction for the pKTH10-containing cells (data not shown).

Differentially expressed genes in the late exponential phase are listed in Table 3. As expected, the highest upregulated genes are *htrA* and *htrB*, encoding the Htr-like proteases. Also moderately elevated transcription levels of the *cssRS* operon, encoding the secretion stress response regulator (CssR) and histidine kinase (CssS), were observed upon AmyQ overproduction, indicating that cells were clearly subjected to secretion stress. In addition, a stimulatory effect was observed for the transcripts of the ribosomal protein RpsB and the genes for the general stress proteins GroEL and DnaK, suggesting that next to secretion stress, also a cytosolic stress, most likely resulting from the presence of misfolded and aggregated proteins, is induced. Moreover, two members of the peroxide stimulon, *ahpF* and *mrgA*, next to the genes involved in cell wall homeostasis (*dltA*, *acpA* and *accC*), showed elevated expression levels. Other genes with known function that were induced upon AmyQ overproduction include *pycA* and *citB* of the TCA cycle and *sodA* and *trxA*, which play an important role in maintaining the redox balance of the cell. The secretion stress directly or indirectly caused increased mRNA levels of four genes with an unknown or predicted function, including *yvfW* encoding a putative iron-sulfur-binding protein and *yvqH* which has similarity to the *E. coli* phage shock protein A that was shown to be induced under anaerobic growth [273] and ethanol stress conditions [193], respectively. Also *ykoJ* and *ydbK* showed increased transcript levels. The former code for a conserved protein of unknown function and the latter is a putative ABC transport system permease protein. Besides the upregulation of a number of genes, some of which have been mentioned above, more than 40 genes were significantly downregulated upon AmyQ overproduction. These genes classify within different functional categories including, amongst others, metabolism of lipids, transport/binding proteins and sporulation (Table 3).

Under natural conditions, most of the secretion takes place during stationary growth phase. Therefore, transcriptome analyses were also performed on AmyQ overproducing and non-overproducing cells in this phase. Again, a strong effect on the transcriptional levels of *htrA*,

htrB, *cssS* and *cssR* was observed (Table 4). Thus, secretion stress also occurs during later stages of growth. From the transcriptional profiling in this growth phase, two main novel findings can be deduced: several genes involved in motility and chemotaxis, including *sigD*, had increased mRNA levels, in contrast to a group of sporulation-related genes, which appeared to have decreased transcription levels (Fig. 1). The list of downregulated sporulation related genes is more extensive, although for some of them the obtained expression ratios were slightly below the statistical cut-offs.

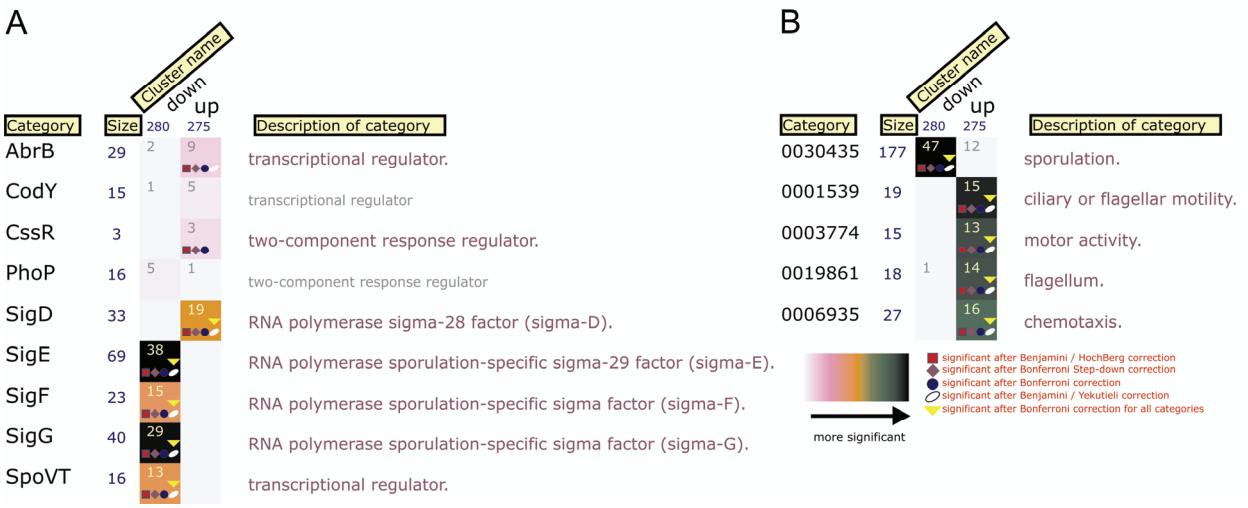


Fig. 1. FIVA analysis of interaction and GO (Gene Ontologies) categories of the transcriptome data.
Genes from DNA-macroarray dataset were divided into up- and down regulated clusters (comparison of gene expression ratios between *B. subtilis* 168 (pKTH10)/*B. subtilis* 168). Panel (A) and panel (B) show "interaction" and GO (Gene Ontology) categories, respectively. The size of each cluster is presented in blue underneath the cluster name. Numbers in each rectangle represent how many genes were up- or downregulated per cluster in each category. The colours of square boxes depict the significance of gene enrichment per category as seen in the legend. Detailed information on significance analyses is available in Blom *et al* [20].

The protein phosphatase encoded by *prpC* (*yloO*) was activated and the product of this gene, together with the PrkC protein kinase, have been shown to be implicated in the sporulation process [158]. Two sporulation genes, *spo0JA* and *spo0E*, had increased transcription levels under the secretion stress but, interestingly, both of them affect sporulation in a negative manner [195]. Expression of several other genes seems also to be affected by the AmyQ overproduction. The products of these genes are associated with different cellular processes such as metabolism of amino acids, lipids and carbohydrates or transport and binding activity (Table 4). For the genes with unknown function, *yloA* and *ykoJ* showed the strongest upregulation (11.2 x and 5.3 x, respectively). Both genes are well conserved among bacterial species. YloA bears similarity to a fibronectin-binding protein as well as to RNA-binding proteins that show homology to the eukaryotic snRNP's. YkoJ was also induced at the earlier time point. This protein contains a putative signal peptide and two PepSY domains which were suggested to have a peptidase regulatory activity in the cell wall proximity [275].

Verification of putative members of the CssRS regulon

As secretion stress clearly affected the expression of the known members of the CssRS regulon, we attempted to identify novel members of this regulon with a focus on the upregulated genes resulting from AmyQ overexpresison. For this purpose we compared the transcriptome profiles of *B. subtilis* 168::*sp* (wild-type) and 168 *cssS*::*sp* (Δ*cssS*), both containing the pKTH10 plasmid. Previous research showed that *htrA* and *htrB* are regulated by the CssRS system and that disruption of *cssS* reduces *htrA* and *htrB* transcription [48,108]. Our transcriptome data confirmed these observations since both genes showed decreased mRNA levels in the *cssS* mutant strain. Beside *htrA* and *htrB,* similar expression patterns from our macroarray approach were observed for *citM* (secondary Mg-citrate transporter), *flgB* operon (flagellar synthesis and chemotaxis), *ylxF*, *yloA* and *ykoJ*.

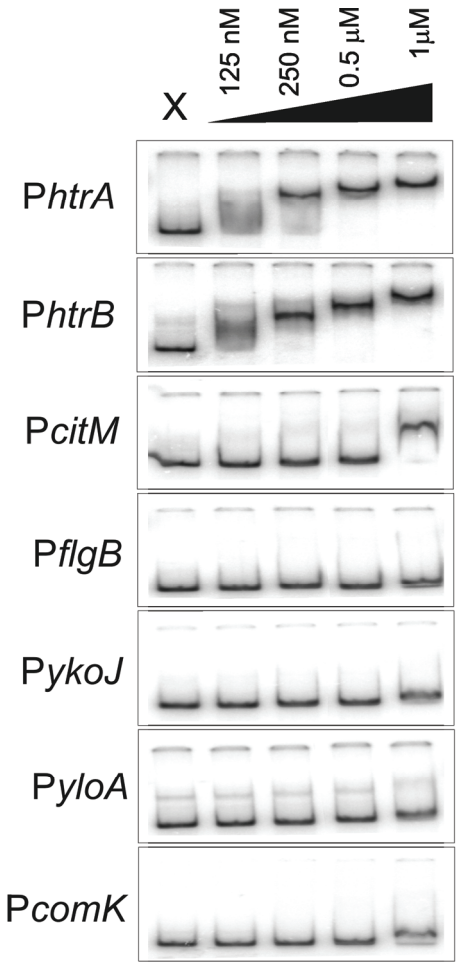


Fig. 2. Binding of CssR-His6 to the indicated promoter regions.
Gel retardation reactions were performed with radiolabeled DNA fragments prepared by PCR and end labeled with ^{32}P. Promoter regions were incubated with increasing concentrations of purified CssR-His6 (see Materials and methods) ranging from 0.125 µM to 1 µM of the protein. In each panel the lane marked with x corresponds to the reaction with no protein added.

Our analysis revealed several genes with altered transcription levels in response to overproduction of AmyQ, in a CssR-dependent manner. Though CssR is a two component regulator, it is unknown if the observed effects are due to direct binding of CssR to the target promoters. In fact, binding of CssR to target promoters has not been reported so far. Therefore, several candidate genes were selected and electrophoretic mobility shift assays

(EMSAs) were carried out, using purified N-terminally His-tagged CssR protein. Phosphorylation of non-tagged CssR by acetyl-phosphate enhances the affinity of the protein for DNA 2-4 fold dependent on the promoter fragment used (E. Darmon, unpublished observations). Figure 2 shows that unphosphorylated, 6xhistidine tagged CssR has a high affinity for the DNA fragments containing the promoters of the *htrA* and *htrB* genes, with an apparent K_D of ~125 nM, therefore, the tagged protein likely mimics the phosphorylated form of the protein. These observations are in good agreement with previous reporter studies that demonstrate the CssR-dependent induction of these genes upon protein overproduction [48,108]. In contrast, unphosphorylated CssR does not to appear to bind strongly to any of the other fragments, with the possible exception of *citM*, for which consistently a reduced mobility was observed at 1μM of CssR protein. We conclude that for the fragments used, only *htrA* and *htrB* contain a high affinity CssR-binding site. As for the other genes, the observed transcriptional changes most likely result from secondary effects.

Sporulation and motility are affected by AmyQ overproduction and secretion

Several genes that were significantly downregulated by AmyQ overexpression belong to the sporulation pathway. This suggests that secretion stress results in inhibition of spore formation. Interestingly, overexpression of a cytoplasmic protein did not affect early sporulation gene expression [118], indicating that secretion stress rather than overproduction stress might be the cause for the observed downregulation. To examine whether the observed downregulation of the sporulation pathway in our transcriptome analysis by AmyQ overproduction was specifically caused by secretion stress, or was a result of an indirect effect caused by AmyQ overproduction, we made use of a sporulation specific reporter strain carrying the *spoIIA* promoter in front of the gene encoding the Green Fluorescent Protein (GFP). The *spoIIA* promoter is directly activated by the key sporulation regulator, Spo0A, and was shown to be a good reporter for cells that initiate sporulation [256]. Plasmids pKTHM10 (AmyQ overexpression), pKTHM102 (AmyQ-Ala overproduction) and pUB110 (empty vector) were introduced in the indicator strain and cells were analyzed for their expression levels by flow cytometry. The pKTHM102 vector is a derivative of pKTHM10 (which is similar to pKTH10) and encodes the AmyQ protein with an Ala-rich signal sequence that renders it inactive in translocation across the cytoplasmic membrane [283]. Importantly, this AmyQ variant does not evoke a secretion stress response [267]. Strains were grown in TY medium and samples were taken at hourly intervals and examined by flow cytometry. As depicted in Fig. 3, strain IIA/Q (P_{spoIIA}-*gfp*, pKTHM10) harbouring an AmyQ overexpression vector with the wild-type signal peptide, was slightly delayed in the activation of the *spoIIA* operon (Fig. 3). A more pronounced effect was observed in cells containing the pKTHM102 plasmid (strain IIA/QAla). Overall, these results show that both the secretion stress and AmyQ overproduction (IIA/QAla) in particular, inhibit the process of sporulation. In addition, this inhibition already occurs at the earliest

stages in spore formation, at the level of phosphorylation of Spo0A, as shown by the single cell analyses.

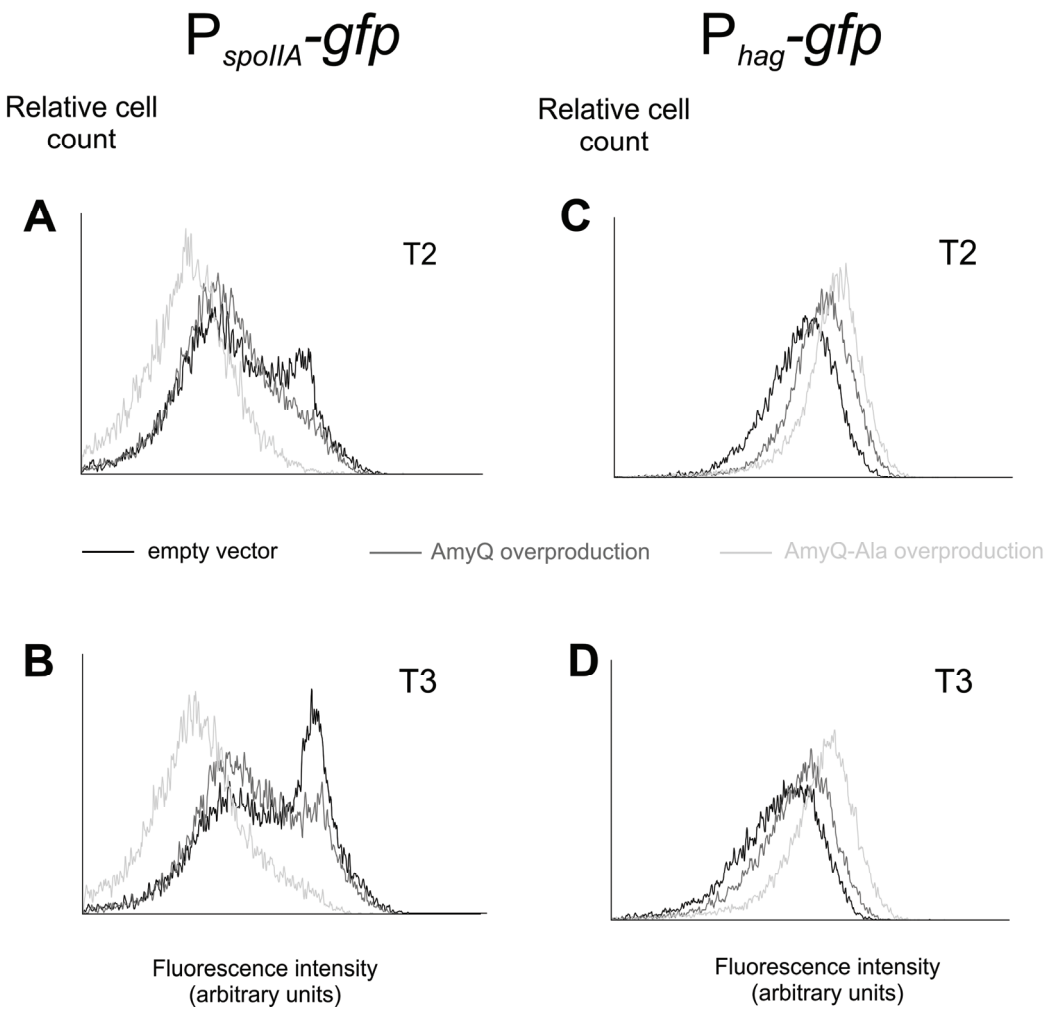


Fig 3. Single cell analysis of sporulation and motility specific gene expression.
Strains were grown in TY medium at 37°C with shaking. Samples for flow cytometric analysis were taken every hour during growth. Two-time points are represented, T2 and T3, which correspond to two and three hours after the entry into stationary phase, respectively. The numbers of cells measured are indicated on the *y* axis and their relative fluorescence is indicated on the *x* axis. (**A-B**) Strains P_{spoIIA}-gfp: IIA/E (empty vector, black line), IIA/Q (AmyQ overproduction, dark grey line) and IIA/QAla (AmyQ-Ala overproduction, light grey line). (**C-D**) Strains P_{hag}-gfp: hag/E (empty vector, black line), hag/Q (AmyQ overproduction, dark grey line) and hag/QAla (AmyQ-Ala overproduction, light grey line).

Another adaptive response that *B. subtilis* utilizes is motility. In this case, cells physically escape from adverse circumstances towards more promising ones to increase their survival chances. Regulation of synthesis of flagella and motility gene expression is known to be orchestrated by an alternative sigma factor, σ^D [96,164]. As described above, we observed a strong upregulation of several genes involved in motility upon AmyQ overproduction. To investigate the increase of flagellar gene expression in *B. subtilis* as a result of protein overproduction, we constructed a strain in which the σ^D-dependent *hag* promoter was fused

to the *gfp* gene (hag-gfp). The *hag* gene codes for the flagellin protein and it was shown before that this gene can be used as a good reporter for studying environmental effects on σ^D-dependent gene expression [169]. The AmyQ overproduction plasmids were introduced in this reporter strain, resulting in strains hag/Q (P_{hag}-*gfp*, pKTHM10), hag/QAla ((P_{hag}-*gfp*, pKTHM102) and hag/E (P_{hag}-*gfp*, pUB110). Cells were grown in TY medium and samples were collected at hourly intervals for flow cytometric analyses. As shown in Fig. 3, the promoter activity of the flagellin gene is the highest for the strain overproducing AmyQ with the Ala-rich signal peptide. This again suggests that the enhanced expression of motility genes is not a direct effect of secretion stress itself, but rather results from AmyQ overproduction.

Effects of *sigD* and *spo0A* deletions on AmyQ overproduction and secretion

The results of the transcriptome analyses have shown that two cellular adaptive processes were clearly affected under the secretion stress conditions. This raised the question whether disruption of these processes would affect heterologous protein secretion by *B. subtilis*.

It has been previously investigated how the yields of protein production can be improved by modification of the components engaged in the late stages of protein secretion, especially the ones influencing the cell wall-associated protease activity or the cell wall composition [268]. To further study whether sporulation and/or motility affect protein secretion in *B. subtilis*, *sigD* (σ^D) and *spo0A* mutations were introduced in a strain containing the pKTHM10 plasmid (AmyQ overproduction).

Mutants of *spo0A* are defective in sporulation since this transcriptional regulator plays a central role in the initiation of this developmental process [195]. The null mutant of σ^D is non-motile and shows reduced levels of autolysins [164]. The production levels of secreted AmyQ in both mutant strains and the parental strain were determined at different time points of growth by means of the EnzCheck Ultra Amylase Assay Kit as described in the Materials and Methods section. In all cases, the amount of active amylase in the culture medium increased from exponential phase (T1) and reached a maximum in the late stationary phase (T4) (Fig 4). However, a clear delay was observed in the *spo0A* mutant, as this strain secreted significantly less amylase than the wild-type strain during exponential phase. In the late stationary phase the difference became minimal between the two strains. More interestingly, under our experimental setup, the *sigD* mutant reached slightly higher levels of the active enzyme at earlier growth phases when compared to the wild-type strain. Again, after four hours upon entry in stationary phase the differences were negligible. These results indicate that when the cells are devoid of the possibility to initiate sporulation or to enter the motile phase, the efficiency and timing of protein secretion by *B. subtilis* are modulated.

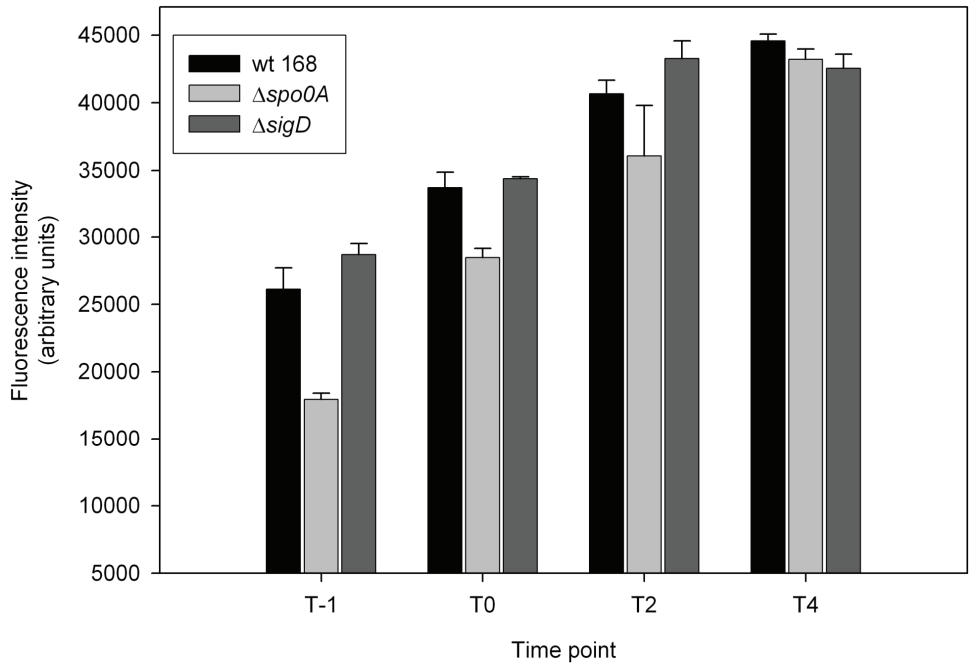


Fig. 4. Activity of secreted amylase from *B. subtilis* 168 wild-type, sporulation- and motility-deficient strains. Amylase activity was quantified in the culture medium of the wild-type strain (black bars), the *spo0A* mutant (light grey bars) and the *sigD* mutant (dark grey bars). All strains were transformed with the pKTHM10 plasmid (AmyQ overproduction) and levels of activity of secreted AmyQ were determined at different time points of growth by using the EnzCheck Ultra Amylase Assay Kit as described in the Materials and Methods section. T0 corresponds to the transition phase of growth. Error bars represent standard errors over three independent biological replicates.

Discussion

Protein secretion, like competence development and sporulation, is one of the post-exponential processes that *B. subtilis* employs as a response to altering growth conditions. The secretion activity is rather low during exponential growth and increases substantially at the onset of stationary phase [200]. It has been shown before that several components of the Sec machinery in *B. subtilis* reach the maximum expression at the end of exponential growth [99], or in the early post-exponential phase [21]. Based on these facts, we set out to study the global cellular response of *B. subtilis* to α-amylase (AmyQ) overproduction and secretion stress during late exponential and stationary phase. In all our macroarray analyses, we observed upregulation of the known targets of the secretion stress regulon, *htrA* and *htrB* (Fig. 1, Table 3 and 4). This validates the quality of the obtained transcriptome data and agrees well with previously published results [48,108]. The comparison of the transcriptional patterns of the control strain and the AmyQ-overproducing strain at the late exponential phase of growth disclosed a rather modest response of *B. subtilis* to the applied stress. This result is in a good accordance with the studies of Hyyrylaïnen *et al*, where the induction of only seven genes upon high-level AmyQ secretion was found [109]. However, our transcriptome analysis at the late exponential phase revealed the activation of a higher

number of genes, which most likely results from the differences in experimental approaches, especially the later sampling in growth in our case. Next to the stimulation of components of the CssRS quality control system, several other genes were weakly induced. Most of these gene products are involved in the stress response evoked by protein misfolding/aggregation events in the cytoplasm, detoxification and fatty acid metabolism, indicating that cells sensed and tried to counteract the adverse conditions.

Our data revealed the CssRS-dependent regulation of several genes, in response to AmyQ overproduction. However, for only two of the (putative) targets (*htrA* and *htrB*), binding of the purified CssR protein to the DNA could be demonstrated, indicating that CssR directly binds to and regulates expression of these operons. How can the apparent CssR-dependent expression of these genes then be explained? Firstly, it is possible that some of the promoter regions tested by EMSA require additional co-factors *in vivo* for a proper binding of CssR. Secondly, phosphorylated CssR might bind different DNA sequences and with other affinities than unphosphorylated CssR. Since our EMSA experiments were carried out with unphosphorylated protein, we cannot exclude that *in vivo* some of the target genes are more strictly dependent on the phosphorylated form of CssR. Another explanation is that the observed effects are indirect and independent of CssR, for instance via increased levels of HtrA, the activity of which was postulated to influence the regulatory effects of a two-component system in *Streptococcus pneumoniae* [222]. Finally, data obtained from a P*ykoJ-gfp* fusion at its native locus, indicates that up-regulation of this gene in response to protein overproduction is indeed CssR-dependent but that promoter activity could not be demonstrated in an area of 500bp upstream of the *ykoJ* open reading frame (reference [186] and unpublished observations). The promoters of *htrA* and *htrB* are remarkably similar, making the *in silico* identification of a CssR binding motif very difficult. However, a Gibbs sampling method [243] identifies a CATTTTTATC motif, that is present in the 300-bp region upstream of both genes. In addition, close matches to this motif (GATTTTTTTC and CATTTTTTTC) can be identified 300-bp upstream of *citM*. The importance of this putative binding sequence remains to be established in future investigations.

The most striking outcome of the transcriptome analysis of stationary phase cells was the influence of protein overproduction on the processes of sporulation and motility. The effect on sporulation was much more pronounced at the later stages of growth and reduced expression levels of many SigE, SigF and SigG regulon members were observed in AmyQ overproducing cells (Fig. 1). Interestingly, the level of *spo0E* transcript (phosphatase of Spo0A~P) in the AmyQ-overexpressing cells is higher than in the wild-type (pUB110) due to which the Spo0A~P pool might be drained. The inhibition of sporulation was confirmed by *gfp*-reporter experiments, and these analyses showed that sporulation is already blocked at the earliest stages (at the level of Spo0A~P) of amylase overproduction (Fig. 3). Using a non-translocated AmyQ variant, the differences between secretion stress and AmyQ production stress were examined. This approach led us to conclude that the impaired sporulation phenotype is most likely not an exclusive and direct effect of secretion stress, but is rather

caused by high-level cellular protein production and accumulation, possibly at the membrane. In fact, overproduction stress seems to inhibit sporulation more efficiently than secretion specific stress (Fig. 3). It is plausible that the unprocessed AmyQ-Ala variant is still targeted to the Sec translocon and thereby leads to a kind of membrane-congestion stress. This type of stress would account for the observed effect on sporulation and motility and would be less severe in case of the wild-type AmyQ variant which is translocated across the cytoplasmic membrane.

Sporulation and protein secretion are multistep- and energy consuming processes and since protein production and secretion take place earlier in the life cycle of *B. subtilis* than endospore formation, it most likely inhibits the latter due to energy constraints. Our data suggests that *B. subtilis* cells sense a variety of environmental and intracellular signals under protein overproduction stress which leads to the decision that commencing the energy-consuming spore development is not appropriate since it is highly unlikely that this process can be completed successfully.

Another survival strategy which *B. subtilis* employs as an adaptive response under unfavourable environmental conditions is motility. Remarkably, our transcriptome profiling revealed that a substantial part of the motility regulon was induced by high-level protein production and secretion. The alternative sigma factor, σ^D, orchestrates the transcription of many genes whose products are in flagellar assembly, motility, chemotaxis and autolysis [96,164]. In a complex medium the amount of the σ^D protein increases during growth, reaching the maximum level at the transition point [169,170]. By single-cell flow cytometric analysis, we show that AmyQ overproduction prolongs the motile phase of *B. subtilis* (Fig. 3). One of the explanations for this prolonged motile phase includes the putative replacement and competition of alternative sigma factors for core RNA polymerase (RNAP) during stationary phase, a phenomenon well-documented for both *E. coli* [66,163] and *B. subtilis* [100,117,151]. The consecutive sporulation-specific sigma factors conduct developmental events during endospore formation [59]. Since sporulation is blocked upon protein overproduction (Fig. 3), SigD does not have to compete with sporulation-specific sigma factors for free RNAP. However, it has been demonstrated that RNAP is present in excess in *B. subtilis* cells and the engagement of anti-sigma factors or other mechanisms, which would give rise to deactivation of σ^A, could be conceived [74]. An extended expression of the motility regulon can provide cells with a rescue machinery under the deficiency of sporulation but, on the other hand, the members of yet uncharacterized σ^D-regulated genes could be also involved in helping cells to survive stress circumstances.

Our transcriptome results, combined with the single cell analyses, predicted that both initiation of sporulation and motility play an important role in combating protein secretion stress. Indeed, when either *spo0A* or *sigD* were mutated, protein secretion was significantly affected (Fig. 4). It is demonstrated that the presence of a functional *spo0A* is required for efficient protein secretion, while the removal of *sigD* moderately but consistently increases secretion (Fig. 4). It has been shown before that a *sigD* mutation causes an increased

extracellular accumulation of an artificial cell wall-binding lipase by an unknown mechanism [128]. Likewise in our experiments, the secretion of enzymatically active AmyQ turned out to be enhanced in the *sigD* mutant, especially in earlier phases of growth, which renders this mutant strain a good host for efficient extracellular protein production (Fig. 4).

Supplementary material

The slide images and raw data mentioned in this manuscript are available from http://molgen.biol.rug.nl/publication/secstress_data/

Acknowledgements

The authors would like to thank Holger Jahr and Jan Jongbloed for their contribution in the initial stages of the project, Elise Darmon and Sacha van Hijum for helpful discussions and Rutger Brouwer for the preparation of the website with supplementary materials. The work of ATL and GB was supported by the grant IGE01018 from the Dutch Ministry of Economic Affairs. WKS and JWV were supported by Grant 811.35.002 and ABC-5587 from the Netherlands Organization of Scientific Research, Earth- and Lifesciences (NWO-ALW) and Technology Foundation (NWO-STW), respectively. Work performed by EJB was supported by grant 050.50.206 from the NWO-BMI.

Table 3. Differentially expressed genes in response to AmyQ overproduction in the late exponential phase of growth.
[1)]Fold represents relative change in expression levels between the stressed and non-stressed cells.

gene	Fold[1]	Function	Functional group
htrA	22.7	serine protease (heat-shock protein)	Adaptation to atypical conditions
htrB	7.8	serine protease (heat-shock protein)	Adaptation to atypical conditions
rpsB	5.0	30S ribosomal protein S2	Protein synthesis; ribosomal proteins
ahpF	2.6	alkyl hydroperoxide reductase (large subunit) and NADH dehydrogenase	Detoxification
citB	2.5	aconitate hydratase (aconitase)	Metabolism of carbohydrates and related molecules; TCA cycle
groEL	2.4	class I heat-shock protein (chaperonin)	Protein folding
acpA	2.3	acyl carrier protein	Metabolism of lipids
ftsZ	2.2	cell-division initiation protein	Cell division
sodA	2.1	superoxide dismutase	Detoxification
cssR	2.0	two-component response regulator involved in the control of cellular responses to protein secretion stress	Sensors (signal transduction)
cssS	1.9	two-component sensor histidine kinase involved in the control of cellular responses to protein secretion stress	Sensors (signal transduction)
accC	1.9	acetyl-CoA carboxylase subunit (biotincarboxylase subunit)	Metabolism of lipids
yerP/swrc	1.9	involved in the efflux of surfactin	Detoxification
mrgA	1.8	DNA-binding stress protein	Adaptation to atypical conditions
DNAK	1.7	class I heat-shock protein (chaperone)	Protein folding
pycA	1.7	pyruvate carboxylase	Metabolism of carbohydrates and related molecules; main glycolytic pathways
dltA	1.5	D-alanyl-D-alanine carrier protein ligase	Cell wall
trxA	1.5	thioredoxin	Membrane bioenergetics
splB	-2.0	spore photoproduct (thymine dimer) lyase	Sporulation
xpaC	-2.0	hydrolysis of 5-bromo 4-chloroindolyl phosphate	Metabolism of phosphate
tenI	-2.1	transcriptional activator of extracellular enzyme genes	RNA synthesis; regulation
nasC	-2.1	assimilatory nitrate reductase (catalytic subunit)	Metabolism of amino acids and related molecules
purE	-2.1	phosphoribosylaminoimidazole carboxylase I	Metabolism of nucleotides and nucleic acids

gene	Fold[1]	Function	Functional group
glpT	-2.2	glycerol-3-phosphate permease	Transport/binding proteins and lipoproteins
bioD	-2.2	dethiobiotin synthetase	Metabolism of coenzymes and prosthetic groups
spoIVFA	-2.2	inhibition of SpoIVFB and hypothesised to stabilize the thermolabile spoIVFB product (stage IV sporulation)	Sporulation
sqhC	-2.3	squalene-hopene cyclase	Metabolism of lipids
ribR	-2.3	monofunctional riboflavin kinase	RNA synthesis; regulation
pps	-2.3	phosphoenolpyruvate synthase	Metabolism of carbohydrates and related molecules; specific pathways
lipB	-2.3	extracellular lipase	Metabolism of lipids
gerM	-2.3	germination (cortex hydrolysis) and sporulation	Germination
nucA/comI	-2.3	membrane-associated nuclease	Metabolism of nucleotides and nucleic acids
spoVFB	-2.4	dipicolinate synthase subunit B	Sporulation
mmgD	-2.4	citrate synthase III	Metabolism of carbohydrates and related molecules; TCA cycle
cwlJ	-2.4	cell wall hydrolase	Cell wall
hisP/mdgS	-2.5	methionine - methionine sulfoxide transport system ATP-binding protein	Transport/binding proteins and lipoproteins
mmgA	-2.5	acetyl-CoA acetyltransferase	Metabolism of lipids
blyA	-2.6	N-acetylmuramoyl-L-alanine amidase, peptidoglycan hydrolase	Phage-related functions
ebrB	-2.6	SMR-type multidrug efflux transporter	Transport/binding proteins and lipoproteins
thyB	-2.7	thymidylate synthase B	Metabolism of nucleotides and nucleic acids
licH	-2.7	6-phospho-beta-glucosidase	Metabolism of carbohydrates and related molecules; specific pathways
hmp	-2.8	flavohemoglobin	Membrane bioenergetics (electron transport chain and ATP synthase)
csfB	-2.9	sigma-F transcribed gene	Miscellaneous
cotJC	-3.0	polypeptide composition of the spore coat	Sporulation
fabHB/yhfB	-3.2	beta-ketoacyl-acyl carrier protein synthase III	Metabolism of lipids
cotE	-3.6	spore coat protein (outer)	Sporulation
spoIIIAH	-3.6	mutants block sporulation after engulfment	Sporulation
bceB/ytsd	-3.8	bacitracine efflux ABC transport system permease	Transport/binding proteins and lipoproteins

gene	Fold[1]	Function	Functional group
Unknown genes			
yvfW	3.2	iron-sulfur-binding protein	Similar to unknown proteins (from other organisms)
yvqH	2.6	similar to phage shock protein	Similar to unknown proteins (from *B. subtilis)*
ykoJ	2.1	conserved protein	Similar to unknown proteins (from *B. subtilis)*
ydbK	2.0	probable ABC transport system permease protein	Similar to unknown proteins (from *B. subtilis)*
ywjD	-2.3	similar to UV-endonuclease	DNA restriction/modification and repair
yhfC	-2.5	conserved membrane protein, probably involved in fatty acid synthesis	No similarity to other proteins
ywdL	-2.5	bacterial regulatory helix-turn-helix protein (LysR family)	No similarity to other proteins
yomC	-2.6	N-acetylmuramoyl-L-alanine amidase, peptidoglycan hydrolase	Cell wall
yxjO	-2.6	probable transcriptional regulator (LysR family)	RNA synthesis; regulation
yckI	-2.8	probable polar amino acid transport system ATP-binding protein	Transport/binding proteins and lipoproteins
yhfU	-2.8	flavohemoglobin	Metabolism of coenzymes and prosthetic groups
ytnJ	-2.8	probable methionine sulfoxide oxidase	Detoxification
yvbA	-2.9	transcriptional regulator (ArsR family)	RNA synthesis; regulation
yhcN	-2.9	inner spore membrane protein	Similar to unknown proteins (from other organisms)
yuiF	-3.0	unknown; putative transporter	Similar to unknown proteins (from other organisms)
ycdA	-3.0	unknown	No similarity to other proteins
yjhA	-3.0	unknown	No similarity to other proteins
yngI	-4.1	long-chain fatty-acid-CoA ligase	Metabolism of lipids
ysxE	-5.2	conserved protein	No similarity to other proteins
yosA	-6.2	SP-beta protein	No similarity to other proteins

Table 4. Differentially expressed genes in response to AmyQ overproduction in the stationary phase of growth.
[1]) Fold represents relative change in expression levels between the stressed and non-stressed cells.

Gene	Fold[1]	Product	Functional group
htrA	8.8	serine protease (heat-shock protein)	Adaptation to atypical conditions
citM	6.1	2-oxoglutarate dehydrogenase (dihydrolipoamide transsuccinylase, E2 subunit)	Transport/binding proteins and lipoproteins
htrB	4.7	serine protease Do (heat-shock protein)	Adaptation to atypical conditions
fliF	4.3	flagellar basal-body M-ring protein	Motility and chemotaxis
flgB	3.8	flagellar basal-body rod protein	Motility and chemotaxis
flgC	3.6	flagellar basal-body rod protein	Motility and chemotaxis
fliJ	3.3	flagellar protein	Motility and chemotaxis
fliG	2.9	flagellar motor switch protein	Motility and chemotaxis
fliZ	2.9	flagellar protein	Motility and chemotaxis
fliE	2.9	flagellar hook-basal body protein	Motility and chemotaxis
fliY	2.8	flagellar motor switch protein	Motility and chemotaxis
PrpC/yloO	2.8	PP2C protein phosphatase	Protein modification
ilvD	2.7	dihydroxy-acid dehydratase	Metabolism of amino acids and related molecules
fliI	2.7	flagellar-specific ATP synthase	Motility and chemotaxis
fliL	2.7	flagellar protein	Motility and chemotaxis
sucC	2.5	succinyl-CoA synthetase (beta subunit)	Metabolism of carbohydrates and related molecules; TCA cycle
serA	2.5	phosphoglycerate dehydrogenase	Metabolism of amino acids and related molecules
alsT	2.5	amino acid carrier protein	Transport/binding proteins and lipoproteins
flgE	2.5	flagellar hook protein	Motility and chemotaxis
soj/spo0JA	2.4	centromere-like function involved in forespore chromosome partitioning / negative regulation of sporulation initiation	Sporulation
fliK	2.4	flagellar hook-length control	Motility and chemotaxis
fliH	2.4	flagellar assembly protein	Motility and chemotaxis
dppA	2.4	D-alanyl-aminopeptidase	Transport/binding proteins and lipoproteins
cheV	2.3	modulation of CheA activity in response to attractants	Motility and chemotaxis
TatAd/yczB	2.3	component of the twin-arginine translocation pathway	Protein secretion
flhA	2.2	flagella-associated protein	Motility and chemotaxis
uxaC/yjmA	2.2	glucuronate isomerase, hexuronate utilization	Metabolism of carbohydrates and related molecules

Gene	Fold[1]	Product	Functional group
sigD	2.1	motility, chemotaxis and autolysis sigma factor	RNA synthesis; initiation
flhP	2.1	flagellar hook-basal body protein	Motility and chemotaxis
spo0E	2.1	negative sporulation regulatory phosphatase	Sporulation
dltA	2.1	D-alanyl-D-alanine carrier protein ligase	Cell wall
rpsP	2.1	ribosomal protein S16	Protein synthesis; ribosomal proteins
mcpB	2.0	methyl-accepting chemotaxis protein	Motility and chemotaxis
smf	2.0	DNA processing Smf protein homolog	DNA packaging and segregation
nprE	2.0	extracellular neutral metalloprotease	Metabolism of amino acids and related molecules
fliD	2.0	flagellar hook-associated protein 2	Motility and chemotaxis
fhuC	2.0	ferrichrome transport system ATP-binding protein	Transport/binding proteins and lipoproteins
flgL	2.0	flagellar hook-associated protein 3	Motility and chemotaxis
motB	2.0	motility protein B	Motility and chemotaxis
htpG	2.0	class III heat-shock protein (molecular chaperone)	Adaptation to atypical conditions
mcpA	2.0	methyl-accepting chemotaxis protein	Motility and chemotaxis
qoxA	2.0	cytochrome aa3 quinol oxidase subunit II	Membrane bioenergetics (electron transport chain and ATP synthase)
qoxB	1.9	cytochrome aa3 quinol oxidase (subunit I)	Membrane bioenergetics (electron transport chain and ATP synthase)
flgK	1.9	flagellar hook-associated protein 1	Motility and chemotaxis
ffh	1.9	signal recognition particle-like (SRP) component	Protein secretion
recA	1.7	multifunctional SOS repair regulator	DNA recombination
degU	1.7	two-component response regulator involved in degradative enzyme and competence regulation	RNA synthesis; regulation
spoVC	-2.0	peptidyl-tRNA hydrolase (stage V sporulation protein C)	Sporulation
acoB	-2.0	acetoin dehydrogenase E1 component (TPP-dependent beta subunit)	Metabolism of carbohydrates and related molecules; specific pathways
opuBB	-2.1	choline ABC transporter (membrane protein)	Transport/binding proteins and lipoproteins
spoIIE	-2.3	serine phosphatase / asymmetric septum formation	Sporulation
spoIVFA	-2.4	inhibitor of SpoIVFB	Sporulation
glgA	-2.5	glycogen synthase	Metabolism of carbohydrates and related molecules; specific pathways
sspF	-2.6	small acid-soluble spore protein	Sporulation
scoB/ yxjE	-2.6	3-oxoacid CoA-transferase subunit B	Metabolism of lipids
spoVM	-2.9	required for normal spore cortex and coat synthesis	Sporulation

Gene	Fold[1]	Product	Functional group
sspD	-4.8	small acid-soluble spore protein	Sporulation
spoIIIC	-5.0	RNA polymerase sporulation-specific sigma factor(sigma-K)	RNA synthesis; initiation
spoIVCB	-5.5	RNA polymerase sporulation mother cell-specific (late) sigma factor	RNA synthesis; initiation
sspL	-5.6	small acid-soluble spore protein (minor)	Sporulation
spoVAA	-6.1	sporulation protein VAA	Sporulation
cotK	-6.2	small acid-soluble spore protein	Sporulation
spoIVB	-9.5	serine peptidase	Sporulation
tlp	-9.7	small acid-soluble spore protein (thioredoxin-like protein)	Membrane bioenergetics (electron transport chain and ATP synthase)
<u>*Unknown genes*</u>			
yloA	11.2	unknown; similar to fibronectin-binding protein	Adaptation to atypical conditions
ykoJ	5.3	conserved protein	Similar to unknown proteins (from *B. subtilis)*
yheN	3.0	similar to endo-1,4-beta-xylanase	Metabolism of carbohydrates and related molecules; specific pathways
ylxF	2.6	conserved protein	No similarity to other proteins
yfmT	2.6	similar to benzaldehyde dehydrogenase	Metabolism of carbohydrates and related molecules; specific pathways
yusS	2.5	similar to 3-oxoacyl- acyl-carrier protein reductase	Metabolism of lipids
yusQ	2.3	similar to 4-oxalocrotonate tautomerase	Metabolism of lipids
yqzH	2.2	function unknown and unique	No similarity to other proteins
yxiL	2.2	function unknown and unique	No similarity to other proteins
yopS	2.2	possible transcriptional regulator, SP-beta protein	Similar to unknown proteins (from *B. subtilis)*
ylxH	2.1	similar to flagellar biosynthesis switch protein	Motility and chemotaxis
yolA	2.1	SP-beta protein	No similarity to other proteins
yybN	2.1	function unknown and unique	No similarity to other proteins
yyzE	2.0	putative PTS glucose-specific enzyme IIA component	Transport/binding proteins and lipoproteins
ymfC	2.0	unknown; similar to transcriptional regulator (GntR family)	RNA synthesis; regulation
yorD	2.0	cold shock and salt stress induced protein	No similarity to other proteins
ycbP	-2.1	conserved membrane protein	Similar to unknown proteins (from *B. subtilis*)
yjmG/exuT	-2.0	hexuronate transporter	Transport/binding proteins and lipoproteins
yjmF	-2.1	similar to D-mannonate oxidoreductase, hexuronate utilization	Metabolism of carbohydrates and related molecules

Gene	Fold[1]	Product	Functional group
yjmE	-2.3	mannonate dehydratase, hexuronate utilization	Metabolism of carbohydrates and related molecules; specific pathways
yhfN	-2.6	putative Zn-dependent protease	Similar to unknown proteins (from other organisms)
yjmC	-2.8	similar to malate dehydrogenase, hexuronate utilization	Metabolism of carbohydrates and related molecules; TCA cycle
yxjC	-3.0	conserved membrane protein	Similar to unknown proteins (from *B. subtilis)*
yoaR	-3.2	unknown; similar to unknown proteins	No similarity to other proteins
yngG	-4.5	similar to hydroxymethylglutaryl-CoA lyase	Metabolism of lipids

Chapter 4

Mild and severe challenges with lactic acid trigger totally different regulatory responses in *Bacillus subtilis*.

Andrzej Lulko, Henk van Noordenburg, Harma Karsens, Girbe Buist, Jan Kok and Oscar Kuipers

ABSTRACT

Weak organic acids are commonly used as food preservatives due to their ability to inhibit bacterial growth. Only limited knowledge at the genetic level is available concerning the mechanisms that prokaryotes use to cope with these compounds. Obtaining a better insight into the mechanisms governing resistance of microorganisms to weak organic acids is of great relevance for the food industry, especially with regard to survival of probiotic- or spoilage organisms. In this study we aimed to unravel the defence mechanisms that *Bacillus subtilis* employs in response to lactic acid.

Growth experiments in rich, lactic acid-containing medium at different pH levels showed that a subtle pH downshift from 5.2 to 5.1 led to a sudden growth inhibition. Such a distinct behaviour was not detected for cells exposed to hydrochloric acid indicating that this phenomenon cannot be attributed to acidification alone. Whole genome transcriptome experiments were carried out at pH 5.5 and 5.0, as under such pH values the growth of *B. subtilis* was either unaffected (mild stress) or completely ceased (severe stress), respectively. The transcriptome data revealed that the initial response at pH 5.5 (i.e. after 3 min) involved a clear induction of genes under control of CymR, a regulator that directs biosynthesis and recycling pathways of sulfur-containing amino acids. In addition, the mRNA levels of several genes encoding proteins with oxidoreductase activities, including NfrA, MsrA, YqjM, were elevated. After 10 min other defence mechanisms were engaged in counteracting the stress factor; these included the induction of several members of the ferric uptake regulator (Fur) operon, the activity of the FoF1 ATPase and upregulation of genes encoding (putative) multidrug resistance transporters, such as *ebrB*, *yhcA* and *yubD*. Exposure to lactic acid likely evokes a cell membrane stress resulting in differential expression of many genes involved in lipid metabolism. An extensive transcriptional reprogramming was observed upon severe lactic acid challenge (pH 5.0). Under these conditions the general stress regulon SigB and the detoxification regulon SigW were clearly triggered. Quantitative real time PCR of nine genes, which were selected on the basis of their distinct differential expression, showed a reasonable qualitative correlation with the transcriptome data. Of the four genes selected for overexpression to test their possible role in providing an improved resistance to lactic acid, only overproduction of the membrane protein YoeB resulted in a faster recovery of growth after severe lactic acid challenge, possibly due to the contribution of this protein to stabilisation of the cell envelope.

We show that, within certain boundaries, low environmental pH has a minor impact on cells and that the undissociated form of the weak organic acid is mainly responsible for the impaired growth of cells. Upon mild lactic acid stress conditions, several defence mechanisms are switched on and their concerted action allows *B. subtilis* to grow unaffectedly. In that respect, the activity of FoF1 ATPase seems to be of crucial importance under our experimental conditions. At the transcriptional level, mild and severe challenges triggered totally different responses. Several genes with unknown function were differentially expressed at both pH conditions, which demonstrates that the cellular response and adaptation of *B. subtilis* to lactic acid challenge is still far from fully understood. Overexpression of the selected proteins did not lead to a significantly improved resistance against lactic acid.

Introduction

In natural environments as well as during industrial processes, microorganisms are often challenged with a variety of unfavourable conditions. The response of *B. subtilis* to diverse stress conditions, such as salt-, oxidative-, acid-, alkaline-, ethanol-, heat- and cold-shock stress as well as exposure to an array of toxic compounds has been thoroughly investigated. It has been shown that a complex adaptation network comprising of the alternative sigma factor SigB, sigma factors of extracytoplasmic function, in addition to several two-component signal transduction systems , orchestrates the responses in *B. subtilis* to stress [101,221]. On the other hand, there is limited knowledge about bacterial defence mechanisms related to (weak) organic acids stress. Only recently, four reports describing the influence of salicylic acid, sorbic acid and lactic acid in *B. subtilis*, *Lactococcus lactis* and *Lactobacillus plantarum*, respectively, tackled this question [57,194,241,289].

It is known that the antimicrobial activity of weak acids is related to the undissociated form of these compounds, which can easily pass through the cell membrane [27]. Upon entry into the cytosol (close to neutral pH), the weak acid molecules dissociate into protons and anions, which are both of potential danger to the cell [42]. The intracellular excess of protons interferes with the pH gradient (ΔpH) across the membrane, causing an impaired cellular bioenergetics since ΔpH, in addition to the membrane potential, contributes to the transmembrane proton motive force (PMF) in bacteria [191]. Uncontrolled cytoplasmic protonation dissipates the PMF and hampers functioning of proteins and nucleic acids, which are crucial for survival of the cell. Moreover, toxicity of organic acids may result from the intracellular accumulation of anions that have deleterious effects on the cell as they increase osmolarity of the cytoplasm and also may influence activity of metabolic enzymes [213]. To deal with the stress caused by a weak acid, bacteria have developed various defence mechanisms and these include induction of proton pumps such as the F_0F_1-ATPases, or other cation transport ATPases, the glutamate decarboxylation / transport system and synthesis of alkali, for example via the urease and arginine deaminase pathways [42,45]. Low pH environments may also induce the expression of several proteins involved in the protection or repair of DNA, as well as other proteins that enable proper functioning of the cell [93]. Other alternative responses to acidic stress are the modification of the cell wall or the alteration of the composition of the cytoplasmic membrane as demonstrated for *Streptococcus mutans* and other bacteria [72].

In the present study, the global gene expression pattern of *B. subtilis* in response to lactic acid exposure was investigated by transcriptome analysis of samples taken at two different stress conditions, i.e. either at pH 5.0 or 5.5. The former was meant to disclose mechanisms of adaptation to a mild lactic acid stress, whereas the latter pH served to screen for candidate genes, overexpression of which could potentially offer an improvement of growth under severe lactic acid challenge.

Materials and methods

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1. *B. subtilis* strains were grown in 100 ml bottles at 37°C under vigorous agitation in 20 ml TY-broth (1% tryptone, 0.5% yeast extract, 1.0% NaCl) or TY-broth with 0.1 M lactic acid titrated to a preferred pH with a 10 M NaOH-solution (TYLA) in an Innova 4000 incubator (New Brunswick Scientific). For natural transformation, *B. subtilis* cells were grown in minimal medium [230]. No antibiotics were added during growth under stress conditions. *L. lactis* was grown at 30°C in M17 broth supplemented with 0.5% (wt/vol) glucose (GM17) as standing cultures or on agar plates containing 1.5% (wt/vol) agar. Plasmid DNA was introduced into *L. lactis* by electrotransformation as described by Leenhouts *et al.* [141].

Table 1: Strains and plasmids used in this study

Strains	***Characteristics and description***	***Reference***
L. lactis MG1363	Plasmid-free strain	[78]
B. subtilis 168	*trpC2*	[6]
B. subtilis ATCC 6633	*Subtilin producer*	#
B. subtilis NZ8900	*trpC2*, *amyE*::*spaRK,* Km^r	[22]
B. subtilis BFS2620	*trpC2 yoeB::pMutin4mcs*	Micado[a]
Plasmids		
pNZ8902	Empty SURE expression vector, Em^r	[22]
pNZ8903	Cloning vector containing the wild-type subtilin inducible promoter (P_{spaS}) fused to *gusA*, Em^r	[22]
pNZ-gspA	pNZ8903 containing *gspA*	This work
pNZ-ykgA	pNZ8903 containing *ykgA*	This work
pNZ-yecA	pNZ8903 containing *yecA*	This work
pNZ-yoeB	pNZ8903 containing *yoeB*	This work

MICADO database (http://locus.jouy.inra.fr/cgi-bin/genmic/madbase/progs/ACCUEIL-MUTANT.pl)
Strain collection of Molecular Genetics group, University of Groningen, the Netherlands
Em^r; erythromycin resistance; Km^r, kanamycin resistance

Growth experiments and determination of ATPase activity

To characterize growth of *B. subtilis* in the presence of lactic acid, an overnight culture of *B. subtilis* was diluted in TY-broth to an optical density at 600 nm (OD_{600}) of 0.05. The culture was incubated at 37°C under continuous shaking (240 rpm), until mid-exponential phase (OD_{600}~0.7). Cells were harvested by centrifugation (1 min 12,000 rpm, at room temperature (RT)). Pellets were resuspended in media with the desired pH (TY or TYLA) in bottles or

Greiner-tubes of 50 ml. For growth experiments in a microplate reader (TECAN, GENios) the resuspended cultures were diluted to OD_{600}~0.3, after which 200 or 300 µl of culture was transferred to wells of a 96 wells microtiter plate. The plate was covered with a gas-permeable sealing membrane ("Breathe Easy", RPI corp.). Growth in the microplate reader was monitored for 18 h at 37°C under continuous shaking (250 rpm).

To examine the effects of *B. subtilis* to other acids, TY medium was supplemented with 0.1 M HCl, acetic acid or citric acid. In all cases the pH was set to pH 5.5 with NaOH.

The role of F_1F_0-ATPase activity in *B. subtilis* at low external pH was investigated by using an ATPase inhibitor, *N,N'*-dicyclohexylcarbodiimide (DCCD; Sigma-Aldrich), added to TY or TYLA at 1 mM final concentration. The mid-exponential cells were harvested by centrifugation for 1 min (12,000 rpm, RT) followed by resuspension in the original volume of fresh TY or TYLA with or without (controls) DCCD.

Transcriptome analyses

For DNA microarray analyses, overnight cultures of *B. subtilis* were diluted to an OD_{600} of 0.05 in bottles containing 20 ml TY-broth and incubated at 37°C shaking at 240 rpm, until mid-exponential phase (OD_{600}~0.7, referred as T0). The cells were harvested by centrifugation for 1 min (12,000 rpm, RT) followed by resuspension in the original volume of fresh TY (pH~7.2) or TY with lactic acid at pH 5.5 (TYLA5.5). Samples for RNA isolation and subsequent DNA microarray analyses were harvested by centrifugation for 1 min (12,000 rpm, 4°C) at 3, 15, 30 and 60 min after resuspension in fresh media and the cells were immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation. Basically, the same procedure was used to sample the cells for DNA microarray analyses of *B. subtilis* inoculated in TY with lactic acid at pH 5.0 (TYLA5.0). As *B. subtilis* does not grow under these conditions, time point T0 for the cells in TY was used as a reference for comparing the culture under the lactic acid stress at pH 5.0 for samples that were taken after 15 and 30 min following resuspension in fresh medium.

Production of DNA-microarrays containing specific oligonucleotides for all 4107 open reading frames (per slide two probes of each ORF were spotted) of *B. subtilis* 168, preparation of total RNA, cDNA synthesis, fluorescent labelling, DNA microarray hybridization, washing and scanning as well as image analysis, data normalization and the statistical analysis of global gene expression were described previously by Lulko *et al.* [154]. For the transcriptional profiling at pH 5.0 and pH 5.5, two DNA microarray experiments were carried out (slides contained two probes of each ORF) per time point of sampling.

Strain construction for induced overexpression

To construct the *B. subtilis* subtilin-inducible plasmids pNZ-gspA, pNZ-ykgA, pNZ-yecA and pNZ-yoeB the complete open reading frames of the corresponding genes were obtained by PCR using the primer pairs listed in Table 2, *B. subtilis* 168 chromosomal DNA as a

template and the Phusion (Finnzymes) or Extensor (Abgene) DNA polymerases, according to the instructions of the suppliers. The PCR products were digested with the restriction enzyme pairs indicated in Table 2, according to the supplier's instructions. Subsequently, the DNA-fragments were ligated into the vector pNZ8903 [22] from which the *gusA* insert was removed with the corresponding restriction enzyme pairs.

Table 2. Primer sequences. Restriction enzyme sites are underlined.

Name	Sequence (5'-3')	Restriction sites
gspA_F	CGGGTGACCACCTTGAGGAAAGATGAAATCATGCATATCG	*Bst*EII
gspA_R	CCCAAGCTTTTATTATTGGTTGATCGCCGGATTCCCAATCGTA	*Hin*dIII
yecA_F	CGGGTGACCTCTATGTTGCCAGAATTGCAGCGCTCTATTACT	*Bst*EII
yecA_R	CCGCTCGAGTGGGGGCCATGAAAGCGGGTTCGGGAAT	*Xho*I
ykgA_F	CATGCCATGGACGTATCCATCCCGAAATCTCAGCAC	*Nco*I
ykgA_R	CCGCTCGAGTTATCAATTTCGATAATGTGATAGCCTC	*Xho*I
yoeB_F	GCGGTGACCGCCATGAAAAAATGTCTTCTATTTCTAAC	*Bste*II
yoeB_R	CCCAAGCTTCCGCTTATACCGGGCGTTTTTCTTA	*Hin*dIII

Ligation-mixtures were used to transform to *L. lactis* MG1363 by electro-transformation using a Gene pulser (Biorad) as described elsewhere [141]. Colonies were selected at 30°C on GM17 agar plates containing erythromycin (5 μl/ml, Boehringer). Resistant transformants were checked by colony-PCR with Taq-polymerase (Invitrogen). Plasmids were isolated from the positive *L. lactis* transformants [141] and checked by digestion with restriction enzymes. Isolated plasmids containing the correct insert were then used to transform to *B. subtilis* 168 NZ8900 [22]. Transformants were selected after overnight incubation at 37°C on TY agar plates containing erythromycin (5 μg/ml). Overexpression constructs were checked by DNA-sequencing (UMCG, Groningen).

Protein expression and polyacrylamide gel electrophoresis

B. subtilis cultures were induced at an OD_{600} of ~0.7 by 1% (vol/vol) of subtilin-containing supernatant of strain ATCC6633 prepared as described before [22]. One, two and three hours after induction, cells were separated from the supernatant by centrifugation (1 min, 14,000 rpm) and prepared for SDS 15% polyacrylamide gel electrophoresis as described previously [155]. After separation, proteins were stained with Coomassie Brilliant Blue and their sizes were estimated with a Precision Plus Protein™ Dual Color Standards-marker (BioRad).

Real-time(RT)-PCR and data analysis

Primers of 15 to 20 nucleotides (Table 1 supplemental materials) were designed with Perlprimer [165] using oligomer sequences of the DNA microarray as a target sequence. Primers were designed to obtain PCR-products of around 100-basepairs in length and were

checked for the presence of secondary structures and possible primer-dimers using a web-based application called Netprimer (http://www.premierbiosoft.com/ netprimer/ netprlaunch). Primers rated with a score above 85 by Netprimer were accepted. cDNA was synthesized from total RNA samples, which were additionally treated with DNAse I (Roche) to remove residual genomic DNA before the reverse transcriptase reaction was initiated. The DNAse-treated RNA was purified with a Roche RNA isolation kit and cDNA was synthesized with Superscript^{tm} III reverse transcriptase (Invitrogen) as described for the DNA microarray experiments. The cDNA was purified with the CyScribe GFX purification kit (Amersham). Concentration of the cDNA was measured in a NanoDrop spectrophotometer (NanoDrop Technologies) and diluted to 1 ng/ul solution, which was then used as a template for the qRT-PCR reactions. The reaction mixture contained 20 µl of master-mix (5 µl 10× PCR buffer, 2.5 µl 50 mM $MgCl_2$, 2.5 µl 2 µg/ml BSA, 1.25 µl 10 mM dNTP, 1.0 µl SYBR Green I mix, 7.55 µl MQ water, 0.2 µl Platinum Taq DNA-polymerase), 20 µl of 10 µM forward and reverse primers and 10 µl of 1 ng/µl cDNA. PCR amplification was initiated at 95°C for 2 min followed by 40 cycles of 95°C for 20s, 60°C for 30s, 72°C for 30s. Next to the PCR reactions of selected genes, samples that did not contain either cDNA template or the Superscript III enzyme were included as controls to detect background contamination or genomic DNA contamination, respectively. The real-time RT-PCR reactions were performed in 96-well plates using an optical ICycler iQ real-time PCR detection system (Biorad). The data obtained was transferred to the Plexor software (https://www.promega.com/techserv/tools/plexor/) to analyze amplification-associated fluorescence of SYBR Green signals and derive threshold-cycles (Ct). Three independent RT PCR amplifications from two independent RT reactions were averaged and the amount of cDNA of a target gene was normalized to the level of the control gene *ywjA* (ΔCt). Thereafter, the differences in expression levels were quantified for each gene by the $2^{-\Delta\Delta Ct}$ method [148].

Results and discussion

Growth and pH measurements following lactic acid and hydrochloric acid stress

To establish a proper experimental setup for the whole-genome transcriptome analyses the characteristics of *B. subtilis* 168 growth in response to lactic acid exposure were first investigated. Mid-exponential phase cells grown in TY medium were transferred to fresh TY supplemented with 100 mM lactic acid (TYLA) and adjusted to different pH values (Fig. 1A). To verify that differences in growth in response to the weak organic acid are not due to the acidification process alone, TY medium adjusted to different pH values with HCl was used as a control (Fig. 1C). The comparison of *B. subtilis* growth in TYLA revealed that the effects of this weak acid are strongly dependent on the pH of the medium (Fig. 1A). There was no difference in growth between the strain cultivated in the control TY medium (initial pH 7.2,

Fig. 1B) and the strains grown in media at pH values down to 5.2. Remarkably, a downshift of the pH to 5.1 caused a strong inhibition of growth, which is manifested as a lag phase of around 360 min following the medium switch. The phenomenon was also observed at lower medium pH at which the cells recovered after 480 min at pH 5.0 and after overnight incubation in case of pH 4.9 (data not shown). Interestingly, the reoccurrence of growth after the applied stress coincided in all cases with the increase of medium pH (Fig. 1B). The control experiments carried out in TY medium containing 100 mM HCl titrated to pH 5.0 showed that this acid has no major impact on growth of *B. subtilis* and only at pH 4.5 a slower growth rate was observed (Fig. 1C). With this strong acid no lag in growth occurred and the pH of the medium started to increase immediately after the acid treatment (Fig. 1D). Therefore, we conclude that the observed growth effects upon lactic acid exposure are specific for lactic acid and cannot be attributed to acidification of the medium alone.

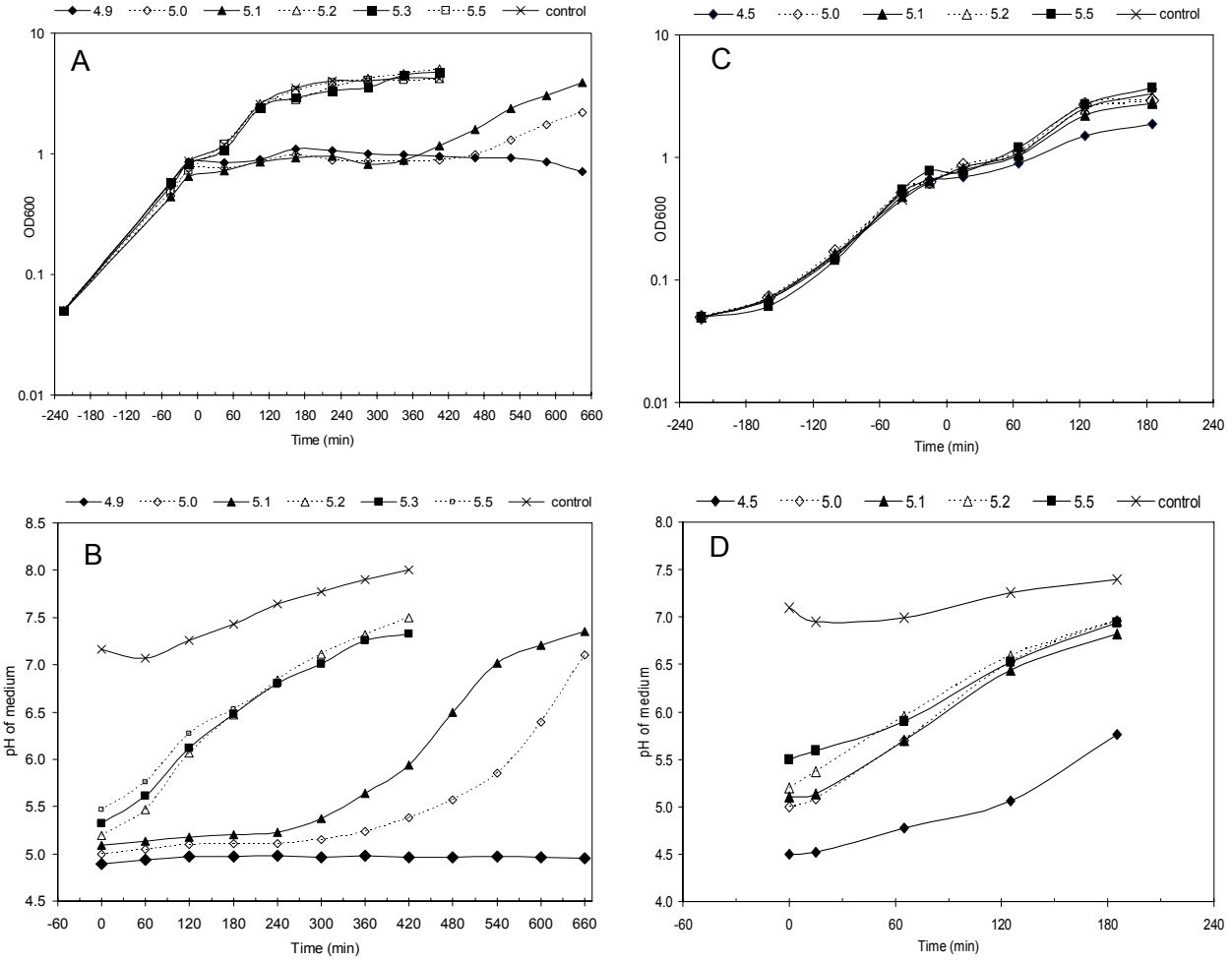


Fig. 1. Growth and medium pH measurements of *B. subtilis* 168 in the presence of lactic acid (panels A and B, respectively) and HCl (panel C and D, respectively) at different pH values. Cells were grown in rich medium until mid-exponential phase after which they were transferred to fresh medium supplemented with either 100 mM lactic acid or 100 mM HCl, adjusted to different pH values as indicated in the accompanying legends (4.9-5.5 for lactic acid or 4.5-5.5 for HCl). The time point of medium replacement is indicated as zero. Control - *B. subtilis* 168 grown in TY medium (pH of 7.2). Results are representative of at least two experiments.

The sensitivity of *B. subtilis* to lactic acid is striking. Under our experimental conditions a pH downshift from 5.2 to 5.1 evoked a dramatic phenotypic response represented by a clear growth inhibition. Based on the Henderson-Hasselbach equation, the calculated concentrations of the undissociated form of lactic acid at pH 5.1 and pH 5.2 were 6.6 mM and 5.3 mM, respectively. As acidification resulted in only minute growth effects, we assumed that the protonated form of the acid at concentrations above 5.3 mM is a major culprit of the growth inhibition. Based on these results, DNA microarray analyses were performed at two different pH values, namely pH 5.5 and 5.0. At pH 5.5 only 2.2% of the lactic acid is non-dissociated, which corresponds to 2.2 mM, a concentration that constitutes around half of the concentration required for the inhibition of *B. subtilis* growth. At pH 5.0, the concentration of protonated lactic acid is three times higher than at pH 5.5, which leads to an arrest in growth. Through applying these two conditions, we were able to monitor what kind of transcriptional reprogramming takes place in *B. subtilis* cells under the mild and harsh lactic acid challenges and based on the latter, select candidate genes for designing strains with an improved resistance against lactic acid.

Regulons affected by lactic acid stress

Global gene expression profiles were monitored at different time points (i.e. 3, 15, 30 and 60 min; referred to as T3, T15, T30 and T60, respectively) after exposure of exponentially growing *B. subtilis* 168 to lactic acid-containing TY medium at pH 5.0 (only two time points, *i*.e. T15 and T30) or pH 5.5 (all four time points). The stressed culture showed an almost identical growth pattern (Fig. 1A) as that of the control one (TY medium, pH~7.2), hence the latter was also sampled for DNA microarray analyses at the same time points. In most of the cases, genes with ≥ two times increased or decreased mRNA levels in at least one of the time point samples (and with a CyberT Bayes p-value < 0.001 or p-value < 0.05) will be discussed here. To gain an overall picture of the transcriptional changes induced by the applied stress conditions, the gene expression ratios from all the time points at both pH's were evaluated with the FIVA software [20] to identify regulons that are influenced by both lactic acid treatments.

The data in Fig. 2 represent a summary of the regulons that passed at least 2 out of 4 multiple testing corrections as performed in FIVA. Only the genes that were significantly differentially expressed (i.e. CyberT Bayes p-value < 0.001 or p-value < 0.05) and that showed at least 1.5-fold differential expression were introduced into the software. To obtain a more complete picture of regulatory interactions the selection criterion is less rigorous than the one described above (2-fold regulation) since it is assumed that co-regulation of a large number of functionally related genes (even with less impressive fold regulation) also offers a possibility to draw sound biological conclusions. Due to many statistical tests performed in FIVA, all the calculations were controlled by means of four multiple testing corrections (for details, see Blom *et al*. [20]).

Severe lactic acid challenge

From the analyses presented in Fig. 2, it can be concluded that there is hardly any overlap between the results at both pH conditions, while in some cases the cells respond oppositely. Lactic acid treatment at pH 5.0 triggered the response of the SigB and SigW regulons at T15 and T30. In both cases, members of these regulons showed higher expression levels as compared to their levels in the non-stressed cells. The effect is opposite to what is observed at pH 5.5. It is well known that the SigB-dependent general stress regulon plays a major role in the *B. subtilis* defense system in adapting to diverse stress conditions [101,221]. However, the role of this alternative sigma factor in response to weak organic acids is rather ambiguous. One study indicated that the general stress response was not triggered in cells encountering sorbic acid, also a *sigB* mutant was not more susceptible to this acid [241,241]. On the other hand, it was shown that salicylic acid and lactic acid activate the general stress response in *B. subtilis* and *Lactobacillus plantarum*, respectively, [57,194]. Our results at pH 5.0 support the latter observation and indicate that the induction of the general stress response depends on the species of bacteria as well as the type of weak organic acid.

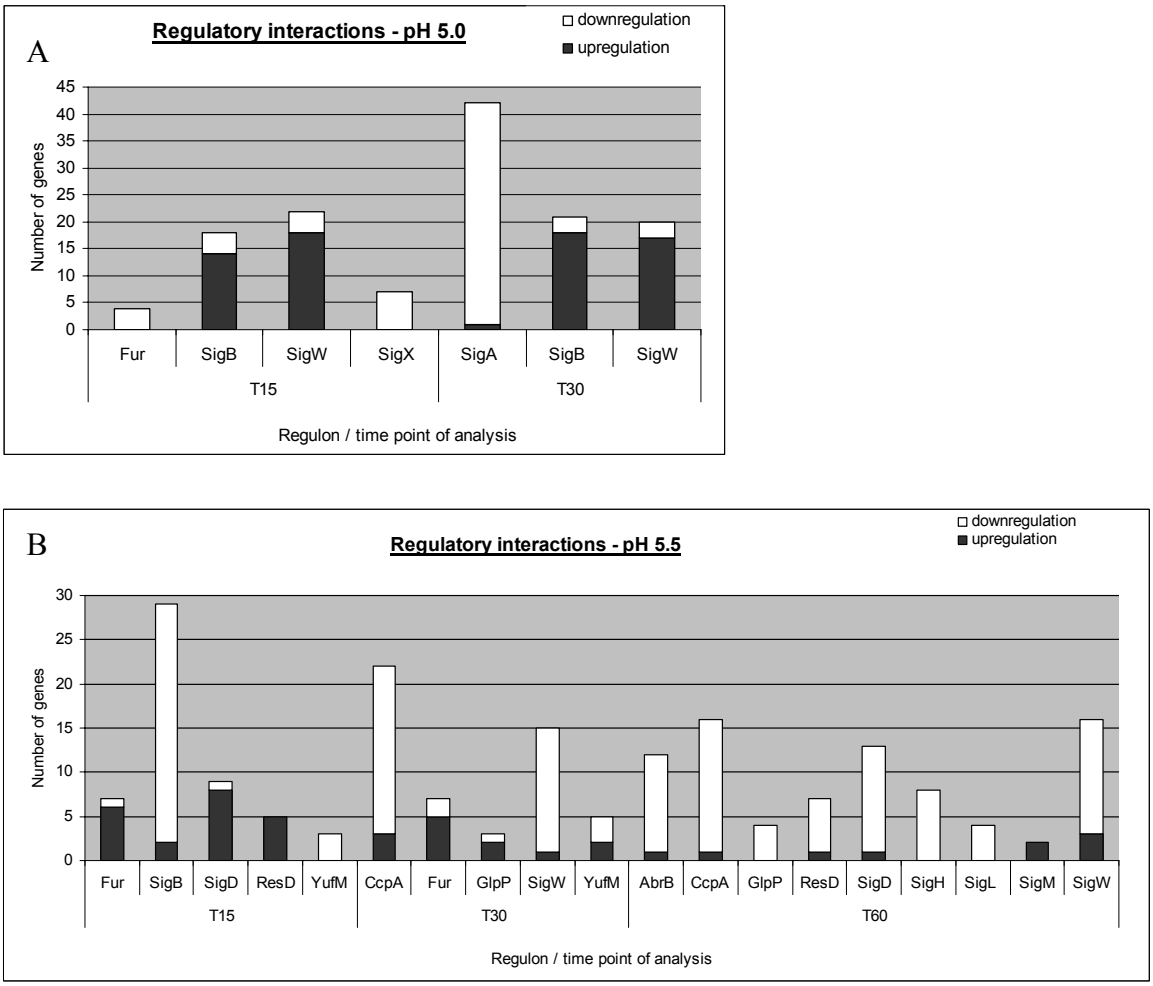


Fig. 2. Overview of regulatory interactions affected by lactic acid at pH 5.0 (panel A) and at pH 5.5 (panel B). Data of regulatory interactions were retrieved from the DBTBS [226].

In addition to the SigB regulon, our data revealed a differential expression of several members of the SigW and SigX regulons. The former was upregulated at pH 5.5, whereas the latter was downregulated. Members of the SigW regulon encode proteins that defend the cell against toxic compounds and are activated by alkaline or salt shock and antibiotics [31,269]. SigX activates expression of genes involved in modifications of the cell envelope through incorporation of positively charged groups [34]. The observed downregulation of the SigX regulon can be possibly explained by the fact that cells under these conditions do not have preference for positively charged groups in cell envelope because this in turn would affect the protonation state of the lactate and would make it more toxic. It has already been suggested that, even though the sigma factors SigW and SigX have overlapping specificity with respect to promoter recognition, they may react to different signals sensed by the cells [166,202,248]. In addition to the responses of the three sigma factors SigB, SigW and SigX at T15, a very strong downregulation was observed at T30 for more than 40 genes controlled by the housekeeping SigA factor. As SigA is a vegetative sigma factor utilized by growing cells, this observation reflects the fact that the presence of excessive lactic acid in the medium causes growth inhibition. Overall, the transcriptomic data at pH 5.0 imply that a severe lactic acid challenge poses a possible threat to cell envelope integrity and bears hallmarks of a toxic effect on the cells, rendering them unable to grow under these conditions.

Mild lactic acid challenge

Further, the data presented in Fig. 2 reveal that not only more regulatory interactions were affected at pH 5.5, but also that the effects on certain regulons (those of SigB and SigW for example) are opposite to those at pH 5.0. The latter notion may indicate that the stress conditions at pH 5.5 were not severe enough to induce the SigB or SigW regulons, although the observed downregulation of these two regulons is difficult to explain at the moment. Even though there was no significant enrichment of genes in any of the regulons for time point T3 at pH 5.5, the expression level of several individual genes was strongly increased at this point in time (more details are presented below). Another interesting observation is the strong effect that the mild lactic acid stress has on the CcpA regulon, but the affected genes did not share a specific type of *cre* sequence. Approximately half of the CcpA regulon genes was repressed at T30 and at T60, which is opposite to the effect seen when *B. subtilis* is treated with either catechol (an aromatic organic compound) or sorbic acid (also a weak organic acid) [237,241,241]. However, in latter studies cells were cultured at a higher pH and media different from the one used here (minimal versus rich medium, respectively). In addition, sorbic acid or catechol- stressed cells showed reduced growth rates whereas in our case cells continued to grow normally after the encountered stress (pH 5.5). It is likely that the effect on the CcpA regulon is caused by considerable changes in the metabolite profile of the cell 30 min after adding lactic acid. As a result, various signal molecules that are able to stimulate a catabolite response may accumulate intracellularly upon lactic acid challenge. It is well known that low molecular weight molecules such as NADP, glucose-6-phosphate

and fructose-1,6-bisphosphate and possibly others can modulate the DNA-binding properties of the CcpA transcriptional regulator [79,125,126,223]. The transcriptome data also suggest that the regulons of ResD (anaerobic response regulator) and SigD (motility and autolysis sigma factor) may play a role in the initial response to the mild lactic acid challenge as several members of these regulons were upregulated at T15. However, this response is transient since at T30 these regulons were not affected while more and at T60 they were even downregulated. In general, short exposure to the stress, such as 3 min, allows monitoring an initial response to the encountered stress factor. Based on the data from Fig. 2, this kind of counter reaction to combat the encountered adverse conditions is rather of a transient nature, as longer exposures (30-60 min) disclose a thoroughly different picture of changes in gene transcription, which represent a long-term adaptation towards lactic acid.

Under adverse conditions the alternative SigB factor drives the expression of many general stress response genes, the products of which provide the cell with a non-specific protection against various stresses [93]. General stress response was not evoked at pH 5.5. On the contrary, half of the SigB regulon was transiently downregulated 15 min after exposure to lactic acid. This indicates that the applied stress conditions were either not severe enough to induce the SigB regulon or that this regulon is not involved in counteracting mild lactic acid stress.

Transcriptome analysis of *B. subtilis* 168 in response to mild lactic acid stress

The experiments carried out at pH 5.5 revealed that changes in gene expression levels occur instantaneously as 18 genes were at least 2 times upregulated already after 3 min of exposure to lactic acid (see Table 4). At this time point there was a clear induction of several genes, including *cysK, ydbM, mtnA, yrhA, yhcL, yxeL, ytlI* and *yrrT*, which all belong to the CymR regulon. Genes under control of CymR are involved in biosynthesis and recycling pathways of sulfur-containing amino acids and available literature data are indicative of a key role of this regulator in the control of cysteine levels in *B. subtilis* [29,63,107]. Also two other genes, *cys*C and *sat*, involved in organic sulfur metabolism were upregulated. The elevated expression of these two genes as well as the ones regulated by CymR likely leads to an increased pool of intracellular cysteine, an amino acid that has been implicated in the oxidative stress response due to its antioxidant properties [147]. In addition, we have observed another remarkable overlap between lactic acid stress and CymR regulation. In a *cymR* mutant decreased mRNA levels of several genes of the SigB regulon as well as the *yoeB* gene were reported [63], which agrees with the expression patterns recorded in our transcriptome analysis. It is tempting to speculate that changes in the cellular cysteine concentration may indirectly influence the genes of the SigB regulon, but further research will be required to validate the role of cysteine in the response to lactic acid stress as well as the connection between this stress and the reduced levels of SigB-regulated genes.

The mRNA levels of the genes of the heat shock-inducible oxidoreductase NfrA and methionine sulfoxide reductase MsrA were also elevated [91,175]. These observations further

indicate that lactic acid treatment impinges on the redox balance of the cells as both proteins have oxidoreductase activities. It is possible that an excessive amount of the lactate anion in the cytoplasm generates oxidizing agents, which could lead to damage of proteins through oxidation of residues such as cysteine and methionine. MsrA could reduce the oxidized methionine and by doing so restore protein function.

In addition, the transcription of *yqjM*, was increased at all time points analysed (except T15, see remark further in the text). YqjM belongs to the Old Yellow Enzyme family of flavoprotein oxidoreductases, the members of which were suggested to be involved in a detoxification defence system in yeast [132]. In *B. subtilis* this protein is activated by toxic xenobiotic compounds as well as by hydrogen peroxide, which led to the conclusion that YqjM activity plays a role in detoxification under oxidative stress conditions [70]. Although the exact function of YqjM remains unclear, it is likely involved in the maintenance of the redox state of the cell as well. All taken together, the results presented above may indicate that lactic acid in the environment leads to immediate toxicity effects and causes an oxidative stress-like response. It seems that the cells try to prevent protein damage caused, for example, by uncontrolled disulfide bond formation that may occur upon excessive cytosolic lactic acid.

The genes of two putative transcriptional regulators of the MarR family, YhbI and YdgJ, showed increased expression levels. Oppositely, the L-lactate permease (LctP) and dehydrogenase (Ldh) encoding genes showed a more than two-fold reduced expression at T3, although they were upregulated at the subsequent time points of analysis. Other downregulated genes at T3 included *ggt* (encoding a gamma-glutamyltranspeptidase involved in glutathione metabolism) and three genes (*ywfF*, *ywcJ* and *yxkJ*) encoding putative transport binding proteins, similar to an efflux protein, and nitrate and metabolite-sodium transporters, respectively. Also the whole *cydABCD* operon required for expression of cytochrome bd was downregulated, although in the case of *cydAB* the statistical criteria were not met. The *cydABCD* operon, together with the already mentioned *lctP*, *ldh* and *ywcJ* genes, belongs to group 2 Fnr-regulated genes for which Fnr mediates the nitrate-dependent repression by anaerobic induction of nitrate reductase production [206]. Since lactate has been identified as one of the major *B. subtilis* anaerobic fermentation products in the growth medium [44], downregulation of these genes indicates that at the very initial stage exposure to lactic acid bears similarities to anaerobiosis.

It has to be noted that in several cases the observed transcriptional effects at T15 were inversed when compared to T3 and T30. At this moment it is difficult to judge whether these effects truly reflect the physiological state of the cells or are caused by technical issues during carrying out of the DNA microarray experiments (*yqjM* is a good example of such an behaviour, see Table 4).

The overall changes in gene expression patterns at the later time points of analysis were much more pronounced, as reflected by more distinctly elevated expression of a higher number of genes. Space limitations do not permit to address all the observed responses.

Since a summary of affected regulons has been presented above, an outline of the most striking and relevant physiological responses will be outlined here. Special focus will be on genes with elevated expression levels as these could be involved in resistance mechanisms. In general, at each time point more than 200 hundred genes showed increased expression levels by at least a factor of two. Among these were several clusters of genes with established functions as well as many unknown genes (Table 4). Nine members of the ferric uptake regulator (Fur) regulon were induced including a.o. the *dhb* operon (involved in the synthesis of the iron siderophore bacillibactin), *fhuD* (part of the ferrichrome ABC transporter system) and the *yku* operon (encoding two flavodoxins). Fur regulates the expression of genes involved in synthesis of transport systems for iron uptake and maintains iron homeostasis in *B. subtilis* [61,206]. Most aerobic bacteria secrete small molecular weight Fe^{3+} chelators (so-called siderophores) that together with their corresponding cell surface transporters form iron supply systems [43]. The Fur protein has been shown to play a role in the acid tolerance of *Salmonella typhimurium* by regulating acid shock genes in an iron-independent manner [85]. The role of Fur, also in *B. subtilis*, extends beyond the regulation of genes involved in iron-sensing and uptake. Apparently, *B. subtilis* somehow experiences iron limitation in the presence of lactic acid and Fur-regulated genes would help to provide the cells with sufficient iron. Excessive lactate anion in the growth medium leads to an increase in the ionic strength of the medium, which could negatively influence the chelating activity of siderophores or their interaction with their cognate membrane receptors. On the other hand, enhanced iron acquisition may represent an adaptive response to environmental stress as the activity of many enzymes that play a role in counteracting low pH or excess of lactate anion rely on iron. This metal, in the form of iron-sulphur clusters or present in heme, functions as a catalytic centre of enzymes involved in various redox processes such as electron transport and metabolism of molecular oxygen. Our transcriptome analyses reveal that several genes involved in heme metabolism and membrane bioenergetics were strongly induced as well. These included *ctaO* (heme O synthase), the *hemAXCDBL* cluster that encodes enzymes required for biosynthesis of uroporphyrinogen III [89] and the *cta* operon required for cytochrome caa3 oxidase biosynthesis. It is interesting to note that Cytochrome caa3 is a heme-containing terminal oxidase which pumps protons across the cytoplasmic membrane as a result of electron transfer of the respiratory chain [271]. Such an activity is crucial under the excessive amount of proton liberated by the dissociation of lactic acid in the cytoplasm, since it serves to regenerate the proton motive force.

Role of the F0F1 ATPase in lactic acid sensitivity

We observed an increase in the expression of the entire operon coding for the subunits of the F0F1 ATP synthase. The involvement of the F0F1 ATPase in pH homeostasis by its capacity to extrude protons from the cells is well documented (reviewed by Cotter & Hill [42]). To investigate a possible role of F0F1 ATPase in the response to lactic acid at pH 5.5, growth of *B. subtilis* was examined in the presence of an inhibitor of the ATPase activity, *N',N'*-

dicyclohexylcarbodiimide (DCCD). A clear inhibition of growth in TYLA medium was seen upon inhibition of ATPase activity, whereas the DCCD-treatment did not affect the cells in a standard TY medium (Fig. 1A). At the same time, the increase of medium pH after imposing a mild lactic acid stress in the presence of DCCD was distinctly slowed down as compared to that of the control culture/not treated with DCCD (Fig. 3B).

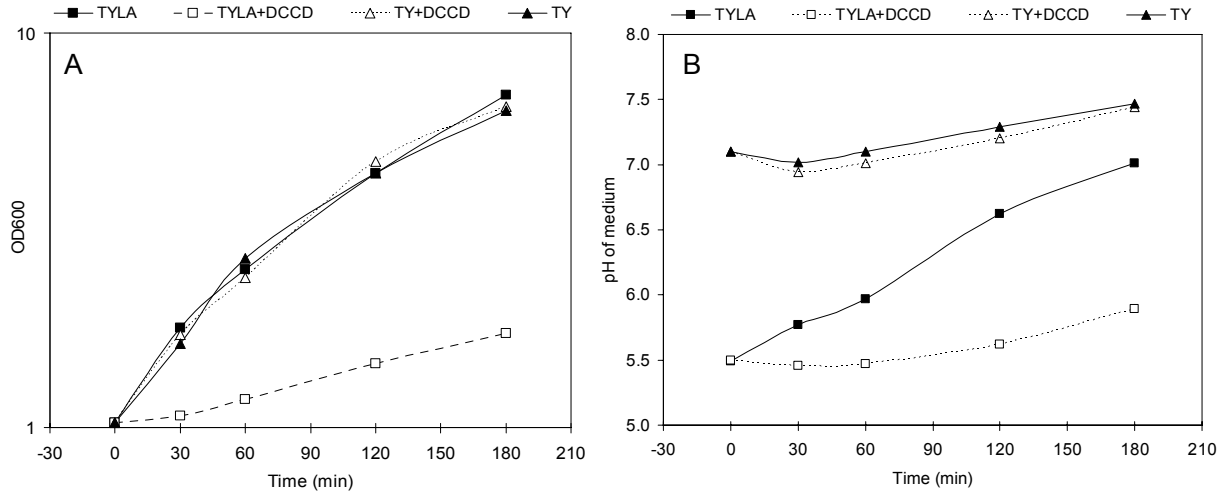


Fig. 3. Role of the F0F1 ATPase in lactic acid sensitivity of *B. subtilis* 168. To measure growth and medium pH (panels A and B, respectively), cultures were first grown in rich medium until the mid-exponential phase and centrifuged after which the cells were transferred to new medium containing either 100 mM lactic acid (TYLA) or 100 mM lactic acid with 1mM *N,N'*-dicyclohexylcarbodiimide (TYLA+DCCD) or standard TY medium (TY; control) as indicated in the accompanying legends. The time point of the media replacement is indicated as zero.

Taken together these results indicate that the prevention of proton translocation by DCCD treatment substantially reduces the growth rate at pH 5.5 and sensitizes *B. subtilis* to the mild lactic acid conditions. However, it has to be noted that resistance to lactic acid does not solely depend on the ATPase activity as growth was not completely inhibited upon DCCD treatment (Fig. 3A). Thus, in addition to a functional F0F1 ATPase complex, also other factors must play a role in the defense mechanisms against weak acid stress. One of the mechanisms involved may lead to alterations of the cytoplasmic membrane in order to adapt to acidification of the environment, for example by changing either the fatty acid composition or the lipid/protein ratio.

Effect of lactic acid on genes related to the cellular membrane

It is known that biological membranes form, together with the peptidoglycan, the first line of defense of bacteria against various stresses. For example, osmotic or cold-shock, both induce extensive modifications in the organization of membrane lipids [214]. Treatment with lactic acid caused differential expression of many genes involved in the metabolism of lipids. Elevated mRNA levels were observed for several members of the fatty acid biosynthetic

pathway (*fab*), an effect which has also previously been observed for *L. lactis* [289] as well as for *B. subtilis*, and *des* encoding the Delta5-lipid desaturase that synthesizes unsaturated fatty acids from saturated phospholipid precursors. The expression of the latter gene is strictly controlled by the two-component signal transduction system DesRK, which has been shown to be essential under cold-shock conditions [2,263]. DesRK activate the *des* gene resulting in a synthesis of unsaturated fatty acids, which also act as negative signaling molecules of *des* transcription. DesRK together with unsaturated fatty acids facilitate survival at low temperatures [14,46]. Interestingly, *desRK* have the same expression pattern as *des*, i.e. a strong induction at 15 min after exposure to lactic acid. The activation of the Des pathway suggests that the cells try to increase the fluidity of the membrane in the presence of the weak organic acid. Previous reports have already implied that changes in membrane lipid composition provide a means of defense against various stresses [2,133,214]. The expected increased membrane fluidity caused by the higher levels of unsaturated fatty acids may influence transport and permeabilization properties of the cytoplasmic membrane and may also prevent membrane damage. Modulation of the lipid-bilayer content could restrict the diffusion of protonated forms of organic acids through the membrane and as such represent a new defense mechanism in *B. subtilis* against lactic acid. It is interesting to note that, in contrast to the induction of *fab* and *des*, the expression of the *bkd* operon (branched-chain fatty acid biosynthesis pathway) was down-regulated. Based on these observations, it is tempting to speculate that elongation (*fab* genes) in combination with desaturation (*des*) of the membrane fatty acids, but not incorporation of branched-chain fatty acids in the membrane, play a crucial role in the adaptation of cells to weak lactic acid stress. Moreover, biological membranes are also protein-rich and these proteins may influence their physical properties. Related to that, the expression levels of several genes that encode proteins with of unknown function but with conserved membrane-spanning domains were increased (data not shown).

Induction of detoxification-related genes

We observed increased mRNA levels of genes of which the products play a role in detoxification (for example, *yerP*) or multidrug resistance (*ebrB, yhcA, yubD*). The *yerP* gene encodes a protein with homology to the RND family (resistance nodulation and cell division) of proton motive force-dependent efflux pumps. YerP may prove to be important under unfavorable excess of lactic acid, as members of the RND family provide bacteria with resistance systems against a variety of harmful compounds [246]. Remarkably, *yhcA* was shown to be highly upregulated by potassium sorbate addition and, unexpectedly, the strain mutated in *yhcA* was highly tolerant to sorbate in minimal medium. However, in rich medium this mutant showed increased vulnerability to sorbate [241,241]. This may indicate that the multidrug resistance transporter confers general resistance to a wider range of weak organic acids. Similarly, the upregulation of genes of some other putative MDR proteins may

play an important role under lactic acid stress, possibly these systems remove the lactic anion or its physiological derivatives out of the cells. The transcript level of the whole *ytrABCDEF* operon was increased. This operon encodes a putative ATP-binding cassette (ABC) transport system involved in acetoin utilization [279], which suggests that under the conditions of lactic acid stress the cells need to metabolize higher amounts of acetoin. However, at the same time the expression level of acetoin anabolism (*alsSD*) and catabolism genes (*acuABC* and *acoABCLR*) was significantly downregulated.

Global transcriptional analysis of *B. subtilis* 168 under severe lactic acid exposure

Global gene expression profiles were monitored at 15 and 30 min after incubation of exponentially growing *B. subtilis* in lactic acid-containing TY medium at pH 5.0. The stressed culture exhibited an immediate growth-arrest for several hours (Fig. 1A), after which cells resumed growing. Thus time point T0 of the non-stressed cells was used for all analysis as a reference. There were as many as 119 and 366 significantly upregulated (ratio ≥2) genes at T15 and T30, respectively. Unexpectedly, FIVA [20] uncovered only a rather limited number of significantly enriched functional categories (see Fig. 2). This can possibly be explained by the simultaneous activation of only a small number of genes of each specific regulon or functional category. Such a broad concerted action of an array of cellular processes would provide a complex defense mechanism in an attempt to survive a sudden environmental change rather than maintain growth (Fig 1). As discussed above, we only detected a moderate induction of the general stress and the detoxification regulons, governed by SigB and SigW, respectively (Fig. 2). A more detailed analysis revealed that 229 genes were differentially expressed (expression ratios above/(below) 3/(-3)) at both time points, of which "y" genes constituted approximately 70% (159 genes, data not shown). This opens promising perspectives to unravel the mechanisms behind severe lactic acid stress, as these genes are good candidates for future research. The genetical reprogramming was more pronounced after 30 min of exposure i.e., higher expression ratios were observed for a greater number of genes. At T30, 367 genes were ≥ 2 upregulated (with an average ratio for the top 10 genes of 19), whereas at T15 it concerned 119 genes (with an average ratio for the top 10 genes of 13.2). The most highly upregulated "y" genes were *yoeB* and *ykgA* at T15 and T30, respectively (Table 5) and the most downregulated ones at both time points were *yybN*, *ykuH*, *ymaA*, *yjzD* (all of unknown function) and *ydjM*. The latter, together with *yoeB*, belongs to the WalRK (previously YycFG) regulon and is predicted to have a role in cell wall metabolism [56]. The *yoeB* and *ykgA* genes, due to their possible role in combating weak acid stress, were selected for the overexpression studies and are discussed further in the text. Other highly induced genes at both time points included *cydAB* (subunits of cytochrome bd ubiquinol oxidase), *gspA* (general stress protein), *ldh-lctP* (lactate dehydrogenase-permease), *dnaX* (DNA polymerase III) and *tyrZ* (tyrosyl-tRNA synthetase). As presented in Table 5, several functional categories were affected including a.o. metabolism of different biomolecules, cell envelope functions as well as adaptation to atypical conditions and

detoxification. In addition, several putative transcriptional regulators showed differential expression patterns, but based on literature searches their possible role in the stress response could not be established. There was almost no overlap between the datasets obtained at pH 5.0 and at pH 5.5 (data not shown) implying that both conditions triggered totally different responses, at least at the transcriptional level. Also at the phenotypical level the cells behaved differently (see Fig. 1), which might partially explain the lack of correlation in the DNA microarray results.

Protein overexpression approach to obtain acid-resistant strains

The DNA-microarray experiments under these conditions were conducted with the main objective to find candidate genes for construction of an overexpression strain with an improved resistance and good growth characteristics under the severe lactic acid challenge. Since at pH 5.0 growth was arrested (Fig. 1), a possible beneficial effect of such strains could be reflected in a form of faster revival of growth, which would not be possible to observe at pH 5.5 due to absence of a lag phase after exposure to lactic acid. The major goal of the experiments performed at pH 5.0 was to screen for strains with improved growth characteristics under the weak acid challenge. To this end, a group of six highly upregulated genes with possible implications in combating the effects of lactic acid stress due to their (predicted) function, in addition to two downregulated genes were selected for real-time quantitative PCR studies in order to validate the DNA-microarray data. The expression ratios from the DNA microarray and qRT-PCR experiments showed good correlations with a tendency of more pronounced relative expression differences measured by qRT-PCR (Table 3). These results confirm the fact that qRT-PCR has a wider dynamic range [47,280]. The confirmation of the DNA microarray results by qRT-PCR, especially for the genes with elevated expression levels, indicates that these could be valid candidates for the generation of overexpression strains with a higher weak organic acid tolerance.

Table 3. Comparison of RT qRT-PCR and DNA-microarrays expression data at pH 5.0. Three independent real time PCR amplifications from two independent RT reactions were averaged (details are described in the Materials and Methods section).

Gene	**T15** qPCR	**T15** DNA arrays	**T30** qPCR	**T30** DNA arrays	Gene function	Effect of overexpression in relation to lactic acid stress
gspA	47.8	15.1	78.4	19.5	general stress protein A	no effect on protection
yoeB	42.4	38.8	30.4	NA	cell wall antibiotic resistance	faster growth recovery
ykgA	12.6	NA	21.3	66.7	similar to arginine deiminase	no effect on protection
cydB	199.8	22.7	260.8	25.7	cytochrome bd ubiquinol oxidase	NA
yuaF	58.9	12.1	44.2	10.9	unknown	NA
ldh	63.8	8.2	40.6	14.4	L-lactate dehydrogenase	NA
yecA	NA	NA	NA	10.9	probable amino acid permease	growth inhibition
yybN	-3.4	-10.1	-9.0	-18.7	unknown	NA
comX	-2.2	-2.3	-7.5	-11.8	competence pheromone precursor	NA

NA – not available/tested

The subtilin-regulated gene expression system (SURE, [22]) was used to obtain several strains overexpressing either a single gene or a whole operon (data not shown). Four constructs, i.e. pNZ-gspA, pNZ-yecA, pNZ-ykgA and pNZ-yoeB were introduced in *B. subtilis* and the strains were tested under various lactic acid stress conditions to examine the contribution of the genes to resistance against lactic acid.

Overexpression of YecA (probable amino acid permease) caused a severe growth inhibition, even when the strain was induced with reduced amounts of subtilin (data not shown). Inconsistent results were obtained with the strain carrying pNZ-ykgA (*ykgA* encodes a putative arginine deaminase) were obtained, *viz*. depending on the experimental replicate, a higher or lower sensitivity to lactic acid was observed for that strain as compared to the empty-vector control (*B. subtilis* pNZ8902; data not shown).

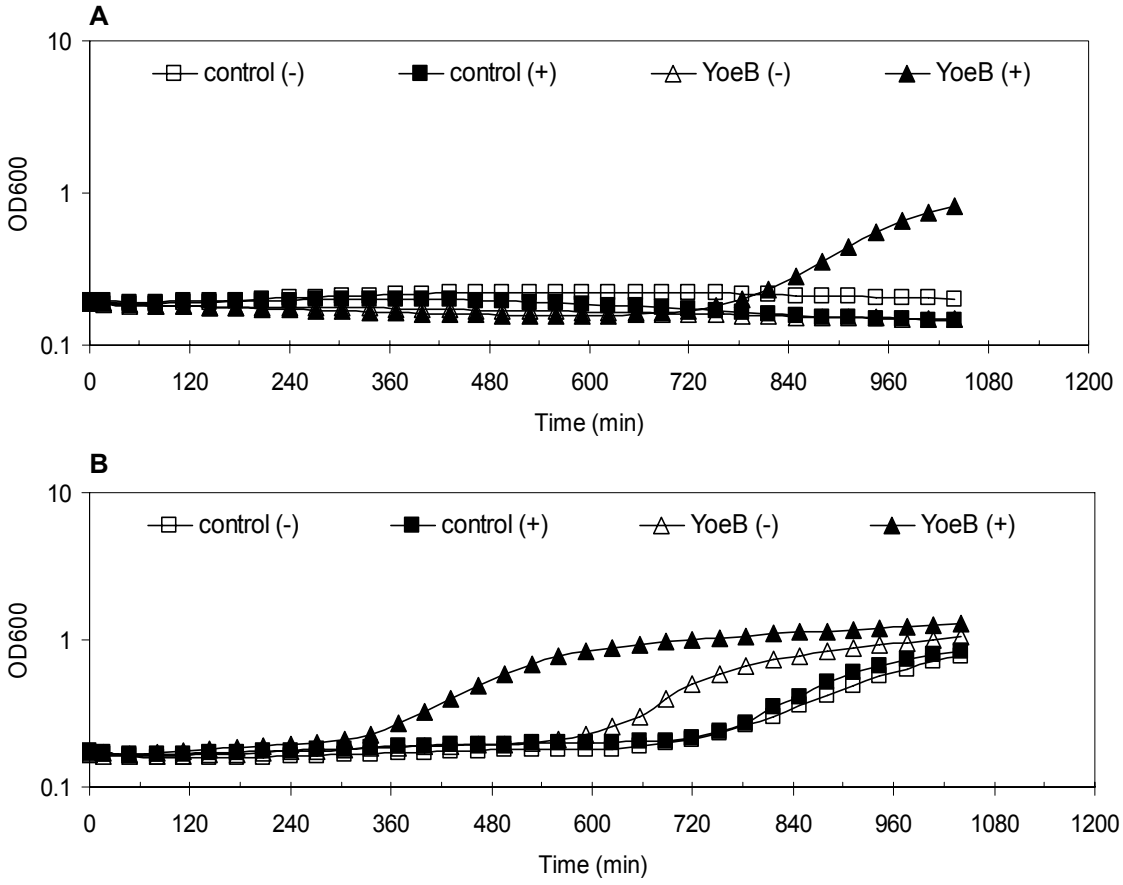


Fig. 4. The YoeB overexpression strain demonstrates accelerated revival after exposure to lactic acid. *B. subtilis* (pNZ-yoeB) was grown first in TY medium in bottles and induced with 1% subtilin at the OD_{600} of 0.6. After approximately 2.5 hours cells were diluted 1:20 in 300 µl of TY medium supplemented with 100 mM lactic acid and adjusted to pH 5.1 (panel A) or pH 5.2 (panel B). Growth was monitored in a TECAN microplate reader. Control- *B. subtilis* 168 strain containing pNZ8902 (empty vector, control). (+) and (-) indicate the presence or absence of subtilin in the medium.

Also overexpression of GspA turned out to have no significant effect on protection against lactic acid. The strain carrying pNZ-yoeB consistently resumed growth after lactic acid challenge at pH 5.2 five to six hours earlier than did the control strain (an example is

represented in Fig. 4). The same holds true at pH 5.1 where the strain carrying pNZ-yoeB was the only strain that revived within 18 h in which the experiment was conducted. As seen in Fig. 4B, also uninduced *B. subtilis* (pNZ-yoeB) started to grow earlier than the control indicating that the SURE system might be slightly leaky, giving a background expression of YoeB leading to an improved response against the encountered unfavourable conditions. Recently Salzberg and Helmann have shown that YoeB is a cell wall-associated, stress response protein that modulates the activity of autolysins [215].

It has been suggested that YoeB inhibits autolytic activity under conditions of nutrient limitation or impaired cell wall synthesis caused by antibiotics treatment, reducing cell death. Under conditions of inhibition of cell wall biosynthesis, *yoeB* was one of the most highly induced genes [34]. Recently, it has also been reported that *yoeB* is under control of WalRK (YycFG), an essential two-component system responsible for the regulation of genes involved in cell wall metabolism of *B. subtilis* [17]. The presence of lactic acid, especially at lower pH, inevitably alters the properties of the cell surface and in that view the role of YoeB would be pivotal as this protein is one of the components of the defence mechanism to restore the outbalanced cell envelope homeostasis.

Concluding remarks and perspectives

The genetic and physiological response of *Bacillus subtilis* to lactic acid is a complex phenomenon. We show that, within certain boundaries, low environmental pH has a minor impact on cells and that the undissociated form of the weak organic acid is mainly responsible for the impaired growth of cells. Subtle changes of pH within an acidic "sensitive" range (in our case pH 5.0-5.5) in the lactic acid-containing medium, even as small as 0.1 pH unit, can lead to a dramatic phenotypic response (as seen in Fig. 1). This indicates that a clearly defined threshold for resistance against the protonated acid exists in *B. subtilis.* Upon mild lactic acid stress, several defence mechanisms are switched on and their concerted action allows *B. subtilis* to grow unaffectedly. In that respect, the activity of F0F1 ATPase seems to be of crucial importance under our experimental setup.
Several genes with unknown function were differentially expressed at both pH 5.0 and 5.5, which shows that the cellular response and adaptation of *B. subtilis* to lactic acid challenge is still largely unknown. Some of these genes may be specifically involved in lactic acid detoxification and deciphering the exact mechanisms behind their action will require further research. Notably a possible role in conferring lactic acid resistance by several of the induced putative multidrug resistance (MDR) proteins or putative ATP-binding cassette (ABC) transporters should be investigated. These kinds of proteins could act as potential efflux pumps for organic acids and thereby prevent their accumulation to toxic intracellular levels.

Overexpression of YoeB resulted in a modestly increased protection against lactic acid. These results may be interpreted in at least two ways. Some of the induced genes may not be directly involved in counteracting harmful concentrations of lactic acid while it is also possible that overexpression of a single protein is not sufficient to confer lactic acid resistance. Instead, a carefully selected set of genes belonging to a weak acid stimulon could be required to achieve a reasonable level of protection against a specific acid challenge.
We expect that the DNA microarray data reported here will not only stimulate further understanding of the poorly defined lactic acid stimulon of *B. subtilis,* but also prove useful in discovering weak acid stress defence mechanisms in other Gram-positive bacteria.

Acknowledgments

The authors would like to thank Evert-Jan Blom for his assistance with the FIVA analysis. This work was supported by IOP genomics grant IGE01018 from the 'Innovatiegerichte Onderzoeksprogramma's' of the Dutch Ministry of Economic Affairs.

Table 4. Differentially expressed genes upon lactic acid treatment at pH 5.5.
B. subtilis genes were divided either in regulons or in metabolic pathways. Only genes of which the expression increased or decreased ≥ 2-fold (with Bayes *p*-value < 0.001 or *p*-value < 0.05) at least in one of the time points are presented.

Gene name	T3	T15	T30	T60	Functional category	Description
CymR regulon / sulfur metabolism						
mtnA	**2.4**	**-3.1**	-	-	2.3 Metabolism of nucleotides and nucleic acids	methylthioadenosine nucleosidase
cysK	**2.1**	**-7.2**	**-3.0**	-2.0	2.2 Metabolism of amino acids and related molecules	cysteine synthetase A
yhcL	1.77	**-4.8**	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to sodium-glutamate symporter
yrhA (mccA)	**3.0**	**-11.2**	-	-	2.2 Metabolism of amino acids and related molecules	O-acetylserine-thiol-lyase
yrhB (mccB)	-	**-6.2**	-	-	2.2 Metabolism of amino acids and related molecules	cystathionine lyase and homocysteine gamma-lyase
yrhC	-	**-4.3**	-	-	6 No similarity	unknown
yrhD	-	**-2.9**	-	-	5.1 Similar to unknown proteins from *B. subtilis*	unknown
yrhE	-	**-2.1**	-	-	1.4 Membrane bioenergetics	unknown; similar to formate dehydrogenase
yrhF	-	**-3.3**	-	**2.1**	5.1 Similar to unknown proteins from *B. subtilis*	unknown; similar to unknown proteins
yrhG	-	**-7.0**	-	**2.2**	1.2 Transport/binding proteins and lipoproteins	unknown; similar to formate transporter
ydbM	-	**-2.1**	1.8	-	2.4 Metabolism of lipids	unknown; similar to butyryl-CoA dehydrogenase
ytlI	**2.1**	-	-	-	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (LysR family)
yrrT	**4.3**	**-13.3**	-	-	6 No similarity	unknown
yxeL	1.9	**-3.0**	-	-	5.1 Similar to unknown proteins from *B. subtilis*	unknown; similar to unknown proteins from *B. subtilis*
cysC	**2.0**	**-12.2**	-	**-2.6**	2.2 Metabolism of amino acids and related molecules	probable adenylylsulfate kinase
msrA	**2.4**	-	-	-	4.2 Detoxification	peptidyl methionine sulfoxide reductase
sat	**2.4**	**-10.5**	-	**-4.2**	2.7 Metabolism of sulfur	probable sulfate adenylyltransferase
Fur regulon						
fur	-	1.6	-1.6	-	3.5.2 Regulation - RNA synthesis	transcriptional repressor of iron uptake
dhbA	-	**7.5**	-	**-2.0**	2.5 Metabolism of coenzymes and prosthetic groups	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase
dhbB	-	**15.6**	-	**-2.6**	2.5 Metabolism of coenzymes and prosthetic groups	isochorismatase
dhbC	-	**9.2**	-	**-2.3**	2.5 Metabolism of coenzymes and prosthetic groups	isochorismate synthase
dhbE	-	**11.9**	-	**-2.7**	2.5 Metabolism of coenzymes and prosthetic groups	2,3-dihydroxybenzoate-AMP ligase
dhbF	-	**15.0**	-	**-2.3**	2.5 Metabolism of coenzymes and prosthetic groups	involved in 2,3-dihydroxybenzoate biosynthesis
fhuB	-	**2.5**	1.6	-	1.2 Transport/binding proteins and lipoproteins	ferrichrome ABC transporter (permease)

Gene name	T3	T15	T30	T60	Functional category	Description
fhuC	-	**4.6**	1.6	-	1.2 Transport/binding proteins and lipoproteins	ferrichrome ABC transporter (ATP-binding protein)
fhuD	-	**2.7**	2.0	-	1.2 Transport/binding proteins and lipoproteins	ferrichrome ABC transporter (ferrichrome-binding protein)
fhuG	-	**3.4**	1.5	-	1.2 Transport/binding proteins and lipoproteins	ferrichrome ABC transporter (permease)
lctP	**-2.1**	-	-	**-4.5**	1.2 Transport/binding proteins and lipoproteins	L-lactate permease
ldh	**-2.7**	**4.1**	-	**-3.0**	1.4 Membrane bioenergetics	L-lactate dehydrogenase
ykuN	-	**5.7**	-	**-2.1**	1.4 Membrane bioenergetics	unknown; similar to flavodoxin
ykuO	-1.6	**18.0**	1.7	**-5.4**	6 No similarity	unknown
ykuP	-	**21.5**	1.8	**-4.6**	1.4 Membrane bioenergetics	unknown; similar to flavodoxin
yxeB	-	**2.4**	1.7	-1.6	1.2 Transport/binding proteins and lipoproteins	unknown; similar to ABC transporter (binding protein)
yusV	-	**2.8**	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to iron(III) dicitrate transport permease
ywbL	-	**2.0**	-	-	5.2 Similar to unknown proteins from other organisms	unknown; similar to unknown proteins
ywbM	-	**2.0**	-	-	5.2 Similar to unknown proteins from other organisms	unknown; similar to unknown proteins
ywbN	-	**2.1**	-	-	5.2 Similar to unknown proteins from other organisms	unknown; similar to unknown proteins
Metabolism of coenzymes and prosthetic groups						
hemA	-	-	1.8	-	2.5 Metabolism of coenzymes and prosthetic groups	glutamyl-tRNA reductase
hemB	-	-	2.0	-	2.5 Metabolism of coenzymes and prosthetic groups	delta-aminolevulinic acid dehydratase
hemC	-	-	**2.2**	-	2.5 Metabolism of coenzymes and prosthetic groups	porphobilinogen deaminase
hemE	-	-	**2.9**	-	2.5 Metabolism of coenzymes and prosthetic groups	uroporphyrinogen III decarboxylase
hemH	-	-	**2.6**	-	2.5 Metabolism of coenzymes and prosthetic groups	ferrochelatase
hemN	-	1.5	-	1.6	2.5 Metabolism of coenzymes and prosthetic groups	coproporphyrinogen III oxidase
hemX	-	-	2.0	-	2.5 Metabolism of coenzymes and prosthetic groups	negative effector of the concentration of HemA
hemY	-	1.6	**4.8**	-	2.5 Metabolism of coenzymes and prosthetic groups	protoporphyrinogen IX and coproporphyrinogen III oxidase
hemZ	1.6	-	-	**2.7**	2.5 Metabolism of coenzymes and prosthetic groups	coproporphyrinogen III oxidase
Membrane functions						
ctaA	-	1.8	**4.6**	-	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (required for biosynthesis)
ctaB	-	**2.4**	1.6	-	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (assembly factor)
ctaC	-	-	**-2.1**	**-7.3**	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (subunit II)
ctaD	-	**5.2**	**-2.0**	**-7.7**	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (subunit I)
ctaE	-	**3.6**	-1.9	**-6.0**	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (subunit III)
ctaF	-	**2.5**	-1.9	**-7.8**	1.4 Membrane bioenergetics	cytochrome caa3 oxidase (subunit IV)

Gene name	T3	T15	T30	T60	Functional category	Description
ctaO	-	**4.7**	-	-	1.4 Membrane bioenergetics	heme O synthase activity
atpA	-	**2.5**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit alpha)
atpB	-	**2.1**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit a)
atpC	-	**2.0**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit epsilon)
atpD	-	**2.5**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit beta)
atpE	-	**2.2**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit c)
atpF	-	1.8	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit b)
atpH	-	**2.0**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit delta)
atpI	-	**2.5**	-	-	1.4 Membrane bioenergetics	ATP synthase (subunit i)
nfrA	**2.0**	**-3.5**	1.6	**2.0**	1.4 Membrane bioenergetics	FMN-containing NADPH-linked nitro/flavin reductase
cydB	-	**5.1**	-	**-3.8**	1.4 Membrane bioenergetics	cytochrome bd ubiquinol oxidase (subunit II)
cydC	**-5.9**	**2.5**	-	-	1.2 Transport/binding proteins and lipoproteins	ABC transporter required for expression of cytochrome bd
cydD	**-6.36**	-	-	-	1.2 Transport/binding proteins and lipoproteins	ABC transporter required for expression of cytochrome bd
Lipid metabolism						
des	-	**4.4**	-1.6	-	2.4 Metabolism of lipids	membrane phospholipid desaturase
desK (yocF)	-	**5.2**	-	**2.3**	1.3 Sensors (signal transduction)	unknown; similar to two-component sensor histidine kinase
desR (yocG)	-1.6	**5.4**	-	-	3.5.2 Regulation - RNA synthesis	unknown; similar to two-component response regulator
fabD	-1.9	**2.9**	1.7	-	2.4 Metabolism of lipids	malonyl CoA-acyl carrier protein transacylase
fabF	-	**2.2**	-	-	2.4 Metabolism of lipids	beta-ketoacyl-acyl carrier protein synthase II
fabG	-1.6	**2.9**	1.6	-	2.4 Metabolism of lipids	beta-ketoacyl-acyl carrier protein reductase
fabHA	-	**2.1**	1.7	**2.0**	2.4 Metabolism of lipids	beta-ketoacyl-acyl carrier protein synthase III
fabHB	-	**6.0**	-	**5.4**	2.4 Metabolism of lipids	beta-ketoacyl-acyl carrier protein synthase III
fabI	-	**2.7**	**2.1**	-	2.4 Metabolism of lipids	enoyl-acyl carrier protein reductase
bkdAA	-	-	**-4.4**	**-3.5**	2.4 Metabolism of lipids	branched-chain alpha-keto acid dehydrogenase E1 subunit
bkdAB	-	-	**-3.9**	**-3.6**	2.4 Metabolism of lipids	branched-chain alpha-keto acid dehydrogenase E1 subunit
bkdB	-	**-2.1**	**-2.8**	**-4.0**	2.4 Metabolism of lipids	branched-chain alpha-keto acid dehydrogenase E2 subunit
bkdR	-	-1.8	**-3.0**	-1.5	3.5.2 Regulation - RNA synthesis	transcriptional activator (control of isoleucine and valine utilization)

Gene name	T3	T15	T30	T60	Functional category	Description
Multidrug resistance						
ebrB	-	-	**2.8**	**2.2**	1.2 Transport/binding proteins and lipoproteins	multidrug resistance protein
yhcA	**2.0**	**-7.6**	**8.8**	**2.3**	1.2 Transport/binding proteins and lipoproteins	unknown; similar to multidrug resistance protein
yubD	-	-	**3.6**	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to multidrug resistance protein
Carbohydrates metabolism						
ytrA	-	1.9	-	1.7	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (GntR family)
ytrB	-	**2.5**	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to ABC transporter (ATP-binding protein)
ytrC	-	**2.3**	-	1.7	2.1.1 Metabolism of carbohydrates and related molecules	unknown; similar to unknown proteins
ytrD	-	**2.0**	-	1.6	2.1.1 Metabolism of carbohydrates and related molecules	unknown; similar to unknown proteins
ytrE	-	**2.7**	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to ABC transporter (ATP-binding protein)
ytrF	-	**2.9**	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to unknown proteins
alsD	-	**-2.6**	**-16.3**	**-7.4**	2.1.1 Metabolism of carbohydrates and related molecules	alpha-acetolactate decarboxylase
alsS	-	-	**-3.4**	-	2.1.1 Metabolism of carbohydrates and related molecules	alpha-acetolactate synthase
acuA	-	-	-1.8	**-2.7**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase
acuB	-	-	-1.9	**-2.8**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase
acuC	-	-	**-2.1**	**-2.7**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase
acoA	-	**-2.1**	-	**-8.2**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase E1 component
acoB	-	**-2.3**	-	**-7.3**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase E1 component
acoC	-	**-2.5**	-	**-3.9**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase E2 component
acoL	-	-	-	**-4.4**	2.1.1 Metabolism of carbohydrates and related molecules	acetoin dehydrogenase E3 component
acoR	-	-	**-3.5**	**-3.3**	3.5.2 Regulation - RNA synthesis	transcriptional activator of the acetoin dehydrogenase operon
yhbI	**2.4**	**-8.7**	**11.5**	**3.9**	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (MarR family)
Other						
yoeB	**-2.2**	-	**-7.6**	-	6 No similarity	unknown;antibiotic-inducible cell wall-associated protein
ggt	**-2.2**	-	-	-	2.5 Metabolism of coenzymes and prosthetic groups	gamma-glutamyltranspeptidase

Gene name	T3	T15	T30	T60	Functional category	Description
ydgJ	**2.57**	-	-	1.8	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (MarR fam.)
yerP	-	**3.6**	-	**2.3**	4.2 Detoxification	unknown; similar to acriflavin resistance protein
yhbI	**2.4**	**-8.7**	**11.5**	**3.9**	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (MarR family)
ydgJ	**2.6**	-	-	1.8	3.5.2 Regulation - RNA synthesis	unknown; similar to transcriptional regulator (MarR family)
y*qjM*	**2.0**	**-4.6**	**2.9**	1.9	1.4 Membrane bioenergetics	unknown; similar to NADH-dependent flavin oxidoreductase
yxkJ	**-4.6**	-	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to metabolite-sodium symport
ywcJ	**-4.8**	-	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to nitrite transporter
ywfF	**-2.2**	-	-	-	1.2 Transport/binding proteins and lipoproteins	unknown; similar to efflux protein

Table 5. Differentially expressed genes upon lactic acid treatment at pH 5.0.
B. subtilis genes were divided in metabolic pathways. Only genes of which the expression increased or decreased ≥ 3-fold (with Bayes *p*-value < 0.001 or *p*-value < 0.05) at least in one of the time points are presented.

Gene/functional category	T15	T30	Description
Membrane bioenergetics and cell wall functions			
cydA	**5.8**	**9.0**	cytochrome bd ubiquinol oxidase (subunit I)
cydB	**22.7**	**25.7**	cytochrome bd ubiquinol oxidase (subunit II)
ldh	**8.2**	**14.4**	L-lactate dehydrogenase
pbpE	**2.3**	**5.1**	penicillin-binding protein
pbpX	-1.8	**-3.4**	penicillin-binding protein
ycnD	**4.4**	-	similar to NADPH-flavin oxidoreductase
ykuN	**-3.4**	**-4.1**	similar to flavodoxin
Transport/binding proteins and lipoproteins			
cydC	-	**11.0**	ABC transporter required for expression of cyt. bd
dppE	-	**5.5**	dipeptide ABC transporter
gabP	1.6	**3.2**	gamma-aminobutyrate permease
lctP	**6.8**	**7.2**	L-lactate permease
mntB	1.5	**3.1**	manganese ABC transporter
mtlA	**2.1**	**3.1**	PTS mannitol-specific enzyme IICBA component
sunT	**-4.6**	**-8.0**	sublancin 168 lantibiotic transporter
ycdH	**-5.0**	**-2.8**	similar to ABC transporter
yceI	1.9	**4.7**	similar to transporter
yceJ	**3.1**	**4.0**	similar to multidrug-efflux transporter
yclN	**-3.7**	**-3.9**	similar to ferrichrome ABC transporter
yclO	**-3.0**	**-2.0**	similar to ferrichrome ABC transporter
yecA	-	**10.9**	similar to amino acid permease
yflA	-	**6.6**	similar to aminoacid carrier protein
yfmD	-	**3.1**	similar to ferrichrome ABC transporter
yhfQ	**-3.5**	-	similar to iron(III) dicitrate-binding protein
yqiH	1.8	**3.2**	similar to lipoprotein
yqiZ	1.9	**3.3**	similar to amino acid ABC transporter
Gene/functional category	T15	T30	Description
yteP	-	**4.2**	similar to transmembrane lipoprotein
ythP	1.6	**3.1**	similar to ABC transporter
yusV	**-3.3**	-	similar to iron(III) dicitrate transport permease
yveA	**2.7**	**5.7**	similar to permease
yvfK	1.9	**3.5**	similar to maltose/maltodextrin-binding protein
ywrB	-	**4.1**	similar to chromate transport protein
yubE	**-17.0**	-	similar to N-acetylmuramoyl-L-alanine amidase
Sporulation			
cotA	-	**3.7**	spore coat protein
cotD	**-3.4**	**-3.2**	spore coat protein
cotP	**-3.2**	**-2.3**	probable spore coat protein
sda	-	**-8.7**	coupling sporulation initiation to replication initiation
seaA	**-2.5**	**-6.5**	involved in spore envelope assembly
spmA	-1.9	**-5.5**	spore maturation protein
spo0B	**-3.3**	**-3.8**	sporulation initiation phosphotransferase
spoIISB	-	**-3.0**	disruption blocks sporulation after septum formation
sspF	-	**3.3**	small acid-soluble spore protein
ypeB	**-4.8**	**-6.5**	similar to unknown proteins

Gene/functional category	T15	T30	Description
Metabolism of carbohydrates and related molecules			
galE	**-2.3**	**-3.7**	UDP-glucose 4-epimerase
mtlD	-	**3.4**	mannitol-1-phosphate dehydrogenase
ycdF	-	**3.1**	similar to glucose 1-dehydrogenase
ycgS	-	**10.2**	similar to aromatic hydrocarbon catabolism
ydaD	-	**3.1**	similar to alcohol dehydrogenase
yfhM	**2.9**	**7.6**	similar to epoxide hydrolase
yolJ	-	**-3.5**	similar to glycosyltransferase
yqjD	-1.7	**-3.1**	similar to propionyl-CoA carboxylase
yveB	1.9	**3.5**	similar to levanase
yvkC	**2.2**	**3.4**	similar to pyruvate,water dikinase
yxbG	-	**5.4**	similar to glucose 1-dehydrogenase
Metabolism of amino acids, nucleic acids, lipids and coenzymes			
aroD	-	**-3.4**	shikimate 5-dehydrogenase
fabHB	-	**3.0**	beta-ketoacyl-acyl carrier protein synthase III
glmS	-1.8	**-3.2**	L-glutamine-D-fructose-6-phosphate amidotransferase
hprT	**-3.2**	**-2.7**	hypoxanthine-guanine phosphoribosyltransferase
ilvD	1.7	**7.0**	dihydroxy-acid dehydratase
mmgA	2.0	**3.0**	acetyl-CoA acetyltransferase
pabC	-1.8	**-3.6**	aminodeoxychorismate lyase
pucD	1.9	**3.1**	xanthine dehydrogenase
purA	**-4.9**	**-6.8**	adenylosuccinate synthetase
racX	-	**4.4**	amino acid racemase
speE	**-4.8**	**-3.3**	spermidine synthase
thiC	-	**3.4**	biosynthesis of the pyrimidine moiety of thiamin
tmk	**-4.1**	-	thymidylate kinase
trpA	**-3.5**	-1.9	tryptophan synthase
yabR	**-3.5**	**-4.5**	similar to polyribonucleotide nucleotidyltransferase
ybcP	**6.0**	1.8	similar to coenzyme PQQ synthesis protein
ycgM	**3.3**	**2.8**	similar to proline oxidase
yhdW	-	**-3.1**	similar to glycerophosphodiester phosphodiesterase
ylbQ	-	**-3.0**	similar to 2-dehydropantoate 2-reductase
ymaA	**-3.7**	**-27.8**	similar to NrdI protein
yqhM	-	**-3.4**	similar to lipoate protein ligase
RNA regulation, synthesis and modification			
abh	-1.7	**-3.1**	transcriptional regulator of transition state genes
gntR	-	**3.6**	transcriptional repressor of the gluconate operon
paiB	1.7	**3.3**	TF of sporulation and degradative enzyme genes
purR	-1.8	**-3.1**	transcriptional repressor of the purine operons
rnpA	**-6.4**	**-7.5**	protein component of ribonuclease P
rpoE	**-2.3**	**-3.5**	RNA polymerase
slr	-	**3.6**	TF of competence development and sporulation
ycgE	**7.2**	**7.2**	similar to transcriptional regulator
ycgK	2.0	**4.3**	similar to transcriptional regulator (LysR family)
yerO	**-2.4**	**-12.5**	similar to transcriptional regulator (TetR/AcrR family)
yhdQ	**-4.9**	**-5.3**	similar to transcriptional regulator (MerR family)
yonR	-	**-3.4**	similar to transcriptional regulator (Xre family)
ytdP	-	**3.4**	similar to transcriptional regulator (AraC/XylS family)
yugI	**-2.1**	**-3.2**	similar to polyribonucleotide nucleotidyltransferase
yxdJ	-	**-3.3**	similar to two-component response regulator [YxdK]

Gene/functional category	T15	T30	Description
yxjO	**-2.9**	**-3.7**	similar to transcriptional regulator (LysR family)
yybE	**-3.0**	-	similar to transcriptional regulator (LysR family)
Protein synthesis, modification and folding			
lepA	-1.6	**-3.4**	GTP-binding protein
bdbA	-2.0	**-5.0**	thiol-disulfide oxidoreductase
dppA	-	**3.8**	D-alanyl-aminopeptidase
yabT	-	**-3.6**	similar to serine/threonine-protein kinase
yfkJ	1.9	**3.0**	similar to protein-tyrosine phosphatase
yvdR	**2.3**	**4.3**	similar to molecular chaperone
Adaptation to atypical conditions and detoxification			
clpC	-	**3.5**	class III stress response-related ATPase
cspC	**-4.7**	**-5.6**	cold-shock protein
ctc	1.8	**3.0**	general stress protein
grpE	-1.5	**-4.2**	heat-shock protein (activation of DnaK)
gsiB	-	**7.1**	general stress protein
gspA	**15.0**	**19.5**	general stress protein
mrgA	**5.1**	-	metalloregulation DNA-binding stress protein
rsbW	**3.2**	**5.9**	negative regulator of sigma-B activity
rsbX	**2.3**	**3.4**	indirect negative regulator of sigma-B activity
sigB	**2.3**	**4.9**	RNA polymerase general stress sigma factor
yceC	**2.3**	**3.2**	similar to tellurium resistance protein
yceE	1.9	**3.0**	similar to tellurium resistance protein
yceF	1.7	**3.4**	similar to tellurium resistance protein
ydaG	**2.5**	**3.4**	similar to general stress protein
ydhE	-	**3.8**	similar to macrolide glycosyltransferase
ydjP	1.8	**6.6**	similar to chloroperoxydase
ykzA (ohrB)	**5.3**	**8.2**	similar to organic hydroperoxide resistance protein
yndN (fosB0	**3.1**	**5.7**	similar to fosfomycin resistance protein
ytxH	1.7	**3.6**	similar to general stress protein
ytxJ	1.7	**3.9**	similar to general stress protein
yxkI	**2.6**	**3.1**	similar to heat-shock protein
Other functions			
comX	**-2.3**	**-11.7**	competence pheromone precursor
divIC	-1.9	**-3.4**	cell-division initiation protein (septum formation)
dnaX	**3.9**	**7.7**	DNA polymerase III (gamma and tau subunits)
flgB	**-2.5**	**-3.7**	flagellar basal-body rod protein
ftsE	**-3.7**	**-3.9**	cell-division ATP-binding protein
gatC	**-2.9**	**-6.2**	glutamyl-tRNA(Gln) amidotransferase
gerAA	-	**3.0**	germination response to L-alanine
gerD	**-2.1**	**-4.7**	germination response to L-alanine
lspA	**-3.5**	**-2.8**	signal peptidase II
prsA	**-2.9**	**-4.4**	protein secretion (molecular chaperone)
recF	**-2.1**	**-3.0**	DNA repair and genetic recombination
recR	**-2.0**	**-3.5**	DNA repair and genetic recombination
tyrZ	**4.8**	**3.4**	tyrosyl-tRNA synthetase (minor)
uvrX	**-4.9**	**-3.6**	UV-damage repair protein
yfhK	**2.3**	**5.4**	similar to cell-division inhibitor
yjoB	1.5	**5.7**	similar to cell-division protein FtsH homolog
ykgA	-	**66.7**	similar to arginine deiminase
yoeB	**38.9**	-	modulates autolysin activity
yuaF	**12.1**	**10.9**	unknown

Chapter 5

Transcriptome analysis of temporal regulation of carbon-metabolism by CcpA in *Bacillus subtilis* reveals additional target genes

Andrzej Lulko, Girbe Buist, Jan Kok and Oscar Kuipers

ABSTRACT

The pleiotropic regulator of carbon metabolism in Gram-positive bacteria, CcpA, regulates gene expression by binding to so-called *cre* elements, which are located either upstream or in promoter regions, or in open reading frames. In this study we compared the transcriptomes of *B. subtilis* 168 and its *ccpA* deletion mutant during growth in glucose-containing rich medium. Although growth was similar, glucose was completely consumed by the wild-type strain in the stationary phase, while it was still present in the culture of the mutant. At that stage direct and indirect effects on gene expression were observed. During exponential growth CcpA mainly influences the carbohydrate and energy metabolism, whereas from transition phase onwards its function expands on a broader range of physiological processes including nucleotide metabolism, cell motility and protein synthesis.
A genome wide search revealed new putative *cre* sites, which could function *in vivo* according to our transcriptome data. Comparison of our data with published transcriptome data of *ccpA* mutant analysis in the exponential growth phase confirmed earlier identified CcpA regulon members. It also allowed identification of potential new CcpA-repressed genes, amongst others *ycgN* and the *ydh* operon. Novel activated members include *opuE* and the *opuAABC*, *yhb* and *man* operons, which all have a putative *cre* site that appears to be dependent on helical topology. A comparative analysis of these genes with the known activated genes *i.e. ackA* and *pta* revealed the presence of a possible upstream activating region (UAR) as has been shown to be functional for the activation of *ackA*. The data suggest that at later growth phases CcpA may regulate gene expression by itself or complexed with other, yet unknown cofactors.

Introduction

The CcpA protein [98,106,174] is a master regulator of catabolite control in many low-GC Gram-positive bacteria [261]. CcpA is a member of the LacI/GalR family of transcriptional regulators [265] which can act as a repressor in carbon catabolite repression (CCR) and as an activator in carbon catabolite activation (CCA) [104,235].

CcpA forms a dimer [219] and its DNA-binding activity is stimulated by complex formation with HPr-Ser-P [52]; [75]] or HPr-like protein Crh-Ser-P [77]. Next to these phosphoproteins also low molecular weight molecules such as NADP, glucose-6-phosphate (G6P) and fructose-1,6-bisphosphate (FBP) can modulate either DNA-binding properties or the interaction with the transcriptional machinery of the CcpA-(HPr-Ser-P) complex [79,125,126,223].

CcpA binds a DNA regulatory sequence known as catabolite-responsive element (*cre*) of which the consensus sequence TGWAANCGNTNWCA (N=any base; W=A or T) was first detected in the promoter region of *amyE* [105,112,173,266]. Later it was discovered that the presence of *cre* either upstream or in the promoter region, or inside a gene does not necessarily imply regulation by CcpA [173]. On basis of *in vivo* operational *cre* sites the following consensus sequence, WWTGNAARCGNWWWCAWW, was proposed. This consensus suggests a high level of degeneracy and indeed, depending on the search method and consensus sequence used, different numbers of *cre* sites can be predicted in the *B. subtilis* genome [50].

The position of the *cre* site(s) relative to the transcription start (TS) of a transcriptional unit determines the regulatory effects upon HPr-Ser-P/CcpA complex binding. In general, promoters with a *cre* site present upstream of the hexameric -35 sequence undergo transcriptional activation as in case of *ackA* and *pta* [198,247]. Interestingly, both genes contain an additional conserved sequence upstream of the CcpA binding site, the presence of which appears to be crucial for transcriptional activation of *ackA* [176]. Recently, a direct positive regulation by CcpA was also demonstrated for the *ilvB* promoter [225,244].

Transcriptional repression by binding of CcpA to a *cre* site downstream of the TS, which blocks elongation by RNA polymerase (RNAP) has been shown. This transcriptional roadblock mechanism has been proposed for the *xyl*, *ara* and *gnt* operons [110,112,281] and for *sigL* [37] and *acsA* [282]. In the latter case affinity of CcpA for the *cre* site was also dependent on the composition of the nucleotides flanking *cre*. Prevention of binding of RNAP to the promoter sequence has been shown for the *acuABC* and *bglPH* operons where a *cre* site overlaps with parts of the promoter region which may prevent transcription initiation [82,135]. Recently, it has been suggested that CcpA does not prevent RNAP from binding to the promoter of *amyE* but that it interacts directly with the RNAP complex already bound to its operator site [127]. In addition, activation or repression by CcpA binding to *cre* is helix-

face dependent since non-integral turns of helix insertions caused relief of CCR of *amyE* [127] or lack of activation as in case of *ackA* [247].

Previously, different transcriptomics technologies have been used, such as Sigma-Genosys macroarrays [19,181] and in-house spotted glass microararrys [276] both containing PCR products of whole open-reading frame (ORF) sequences or commercially available Affymetrix chips consisting of multiple 25-mer oligonucleotides per ORF [150]. However, the common characteristic of these analyses was that they were carried out on mid-exponentially growing cells. The aim of this research was to define the CcpA regulon in time. To this end, we performed whole transcriptome analysis of a *ccpA* mutant strain at four different stages of growth in rich glucose-containing medium and compared it to the well-defined CcpA regulon in mid-logarithmic phase described in literature. We conclude that the CcpA regulon dynamically develops in time and that the impact of the *ccpA* mutation spreads out and intensifies during growth, affecting various cellular processes in growing *B. subtilis* cells. The approach allowed us to further dissect the CcpA regulon over time and permitted identification of new (putative) CcpA targets.

Materials ans methods

Bacterial strains, growth conditions and preparation of cells for RNA isolation

Chromosomal DNA of the *B. subtilis* QB5407 *ccpA* mutant strain [65] was isolated and transformed to *B. subtilis* 168 to ensure the same genetic background of the wild-type and the *ccpA* mutant. Cultures were grown in 250 ml Erlenmeyer flasks by inoculation with "synchronized" cells at OD_{600} 0.01 in the 25 ml TY medium supplemented with 0.5% glucose. Synchronized stocks for inoculation were prepared by diluting overnight (ON) cultures to OD_{600} 0.01 in fresh medium and growing the cells to the mid-exponential phase and diluting again to OD_{600} 0.01. After three cycles of growing to mid-exponential phase synchronized stocks were frozen in glycerol at -80°C for future use. All growth experiments were performed at 37°C with 250 rpm shaking in an Innova 4000 incubator (New Brunswick Scientific). The reason synchronization was to avoid variable lag-phases of the *ccpA* mutant, which were observed when direct inoculation from ON cultures was attempted. Cells for RNA isolation and subsequent microarray analysis were harvested by centrifugation (1 min, 12,000 rpm, 4°C) of early-, mid- and late-exponential phase as well as stationary phase cells corresponding to approximate OD_{600} values of 0.12, 0.8, 3.4 and 5.5, respectively. Cell pellets were immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation.

Design and production of DNA microarrays

The oligonucleotide design was based on the coding sequences of the GenBank RefSeq file NC_000964 for *B. subtilis* 168. The design was performed with OligoArray 2.1-software [211] with the following parameters: oligonucleotide length 70-72 bps, Tm 88-93°C, secondary

structure melting-point 65°C and GC content 42-52%. In the cases where OligoArray 2.1 could not design a suitable oligo, Picky (Complex Computation Laboratory, Iowa state university) was used with less stringent parameters. For 39 ORFs no unique oligo could be designed due to gene duplication within the *B. subtilis* genome and therefore, when possible, oligos were designed manually. For highly homologous paralogues only one oligo was designed. Next to sequences representing the ORFs, additional sequences of antibiotic markers, plasmids and *gfp*/*yfp*/*cfp* probes were spotted (supplementary material). All 4150 sequences were spotted in duplicate on aminosilane glass slides (ez-rays™, Matrix Technologies Corp.). Slide spotting, slide treatment after spotting and slide quality control were performed as described previously [251].

DNA microarray analysis

The general approach for DNA microarray experiments was implemented as described previously [136,251]. Total RNA was isolated with the High Pure RNA Isolation Kit (Roche Applied Science) according to the manufacturer's instructions. RNA quantity and quality were assessed with a Nano Drop ND-1000 spectrophotometer (NanoDrop Technologies) and an Agilent Bioanalyzer 2100 with RNA 6000 LabChips (Agilent Technologies Netherlands BV).

cDNA was synthesized with the Superscript III Reverse Transcriptase kit (Invitrogen) using 25 μg of total RNA as template and 400U of SuperScript™ III RT. The reaction contained 0.5 mM dATP, dCTP, dGTP, 0.3 mM dTTP and 0.2 mM of amino allyl-modified dUTP. The reaction was incubated for 16 hrs at 42°C. The amino allyl-modified cDNA was purified with the Cyscribe GFX purification kit (Amersham Biosciences) using 80% ethanol as wash buffer and 0.1M sodium carbonate solution (Sigma-Aldrich) pH 9.0 as elution buffer. The purified cDNA was labelled with Cy3- or Cy5 mono-reactive Dye (Amersham Biosciences) and incubated at room temperature in the dark for 90 min. Labelled cDNA was purified with the CyScribe GFX purification kit again as described by manufacturer. Dye incorporation and cDNA concentration were assessed with a Nano Drop ND-1000 spectrophotometer (NanoDrop Technologies). The labelled cDNA was hybridized to oligonucleotide microarrays in Ambion Slidehyb #1 buffer (Ambion Europe Ltd.) at 45°C for 18-20 hrs. After hybridization, slides were washed for 5 min. at 37°C in 2 x SSC with 0.5% SDS and 2 x 5 min. at 37°C in 1 x SSC with 0.25% SDS, dipped five times in 1 x SSC 0.1% SDS and then dried by centrifugation (2 min, 2,000 rpm).

The microarrays were scanned with a GeneTac LS V confocal laser scanner (Genomic solution Ltd.). Determination of the individual intensities of each spot was done with ArrayPro 4.5 (Media Cybernetics inc., Silver Spring, MD) with a local corners background correction method and the resulting expression levels were processed and normalized (Lowess method) with MicroPrep [253]. The ln-transformed ratios of the expression levels of *ccpA* mutant versus wild-type were subject to a *t*-test using the Cyber-T tool [149]. For each time point four independent biological replicates (four growth curves with four slides each)

were performed resulting in eight measurements per gene, since each slide contains two duplicate spots for all genes. Genes were selected which had an expression ratio Δ*ccpA*/wt ≥ ±2 at least in one of the investigated time points (values below one were converted by the formula [-1]×[1/ratio value]) and a Cyber-T (Bayes) *p*-value below 0.01. For these genes the values of three time points were added when the expression ratio Δ*ccpA*/wt was at least 1.5 times and had a *p*-value below 0.05. If these criteria were not met the expression ratio was set to one. The final gene list selected in this manner is presented in Table S1 (supplementary material). The data was further analyzed with FIVA (Functional Information Viewer and Analyzer; Blom et al. submitted) to identify overrepresented functional categories in clusters of up and down-regulated genes per individual time point. Two programs were employed for clustering of the time series data; Genesis [236] was used for hierarchical clustering of known members of the CcpA regulon and STEM (Short Time-series Expression Miner; [60]). All the clustering analyses were performed on the $\log_2$ transformed gene expression ratios.

cre site prediction

A weight matrix (Fig. S1) was generated in Genome 2D [10] based on all known *cre's* from DBTBS (**D**ata**B**ase of **T**ranscriptional Regulation in ***B**acillus **s**ubtilis*; [162]) and thereafter it was fed into the MotifLocator tool [1] to search the whole genome of *B. subtilis* for the potential CcpA binding sites. This approach yielded 311 *cre's* with a score above 0.85 (Table S1). The position of *cre* site was always related to a known or deduced TS site of a gene.

Glucose utilisation

D-glucose depletion from medium during growth was determined for the wild-type and the *ccpA* mutant strain by means of the D-Glucose enzymatic bioanalysis/food analysis kit (R-biopharm). Samples were collected at one hour intervals from the inoculation point till eight hours of growth. Supernatants were separated from cells by centrifugation and kept at 80°C for 15 min and the D-glucose concentration was measured according to the manufacturer's instructions.

Results

The *B. subtilis* CcpA modulon in glucose-rich medium was analyzed at different stages of growth by comparing the transcriptional profiles of *B. subtilis* 168 cells with its isogenic *ΔccpA* strain. Samples for DNA-microarray analysis were collected in early (E), middle (M) and late exponential (transition; T) phase, as well as three hours upon entry into stationary phase (S). The difference in growth rate between the strains was small (Fig. 1), which allowed a relatively easy estimation of appropriate sampling times.

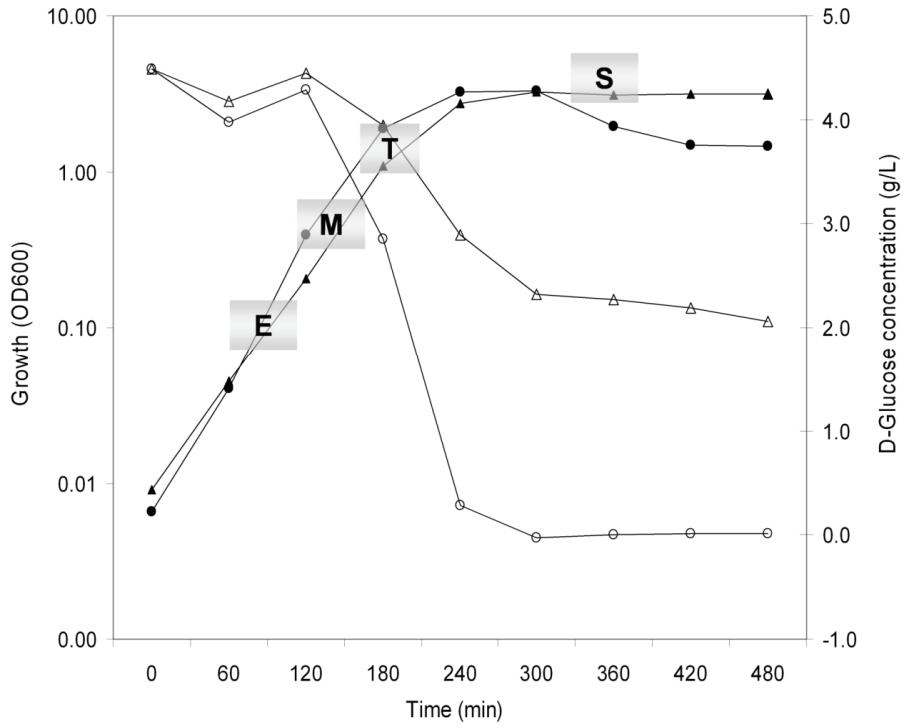


Fig. 1. Growth (closed symbols) and glucose (open symbols) utilisation (in g/L) of *B. subtilis* 168 wild-type strain (circles) and *ccpA* mutant strain (triangles) in TY with 0.5% glucose. Cultures were grown at 37°C under vigorous shaking (250 rpm). Samples for the microarray analysis were collected in early (E), middle (M) and late exponential (transition) phase (T) as well as in stationary phase (S).

Global impact of the *ccpA* mutation at various stages of *B. subtilis* growth

To identify differentially expressed genes between the wild-type and the *ccpA* mutant we subjected the transcriptome data in the form of Lowess-normalized expression ratios [253] to the Cyber-T statistical package [149]. This program was used to evaluate the significance of the Δ*ccpA*/wt expression ratios. The gene expression ratios from the four individual time points was analyzed by the FIVA software (Functional Information Viewer and Analyzer; Blom et al., submitted) to identify which general functional classes (categories) are affected upon *ccpA* mutation and to distinguish the temporal aspect of these changes. Examination of clusters of orthologous groups (COG) of proteins [240]showed that the effects of the mutation in the early exponential phase concern exclusively 'carbohydrate transport and metabolism' as well as 'general energy production and conversion' processes (Fig. 2). The majority of these genes are up-regulated in the *ccpA* mutant. The same two COG categories become even more strongly influenced by CcpA throughout the exponential phase. In the stationary phase gene regulation in the 'energy production and conversion' category shows a swap from up- to down-regulation. The transcriptional profile alters in time since already in the mid-exponential phase additional COG's are significantly influenced by CcpA, namely 'cell motility and secretion' as well as 'nucleotide and amino acid transport and metabolism'. When the clusters of the mid-logarithmic phase are compared to the clusters of the transition

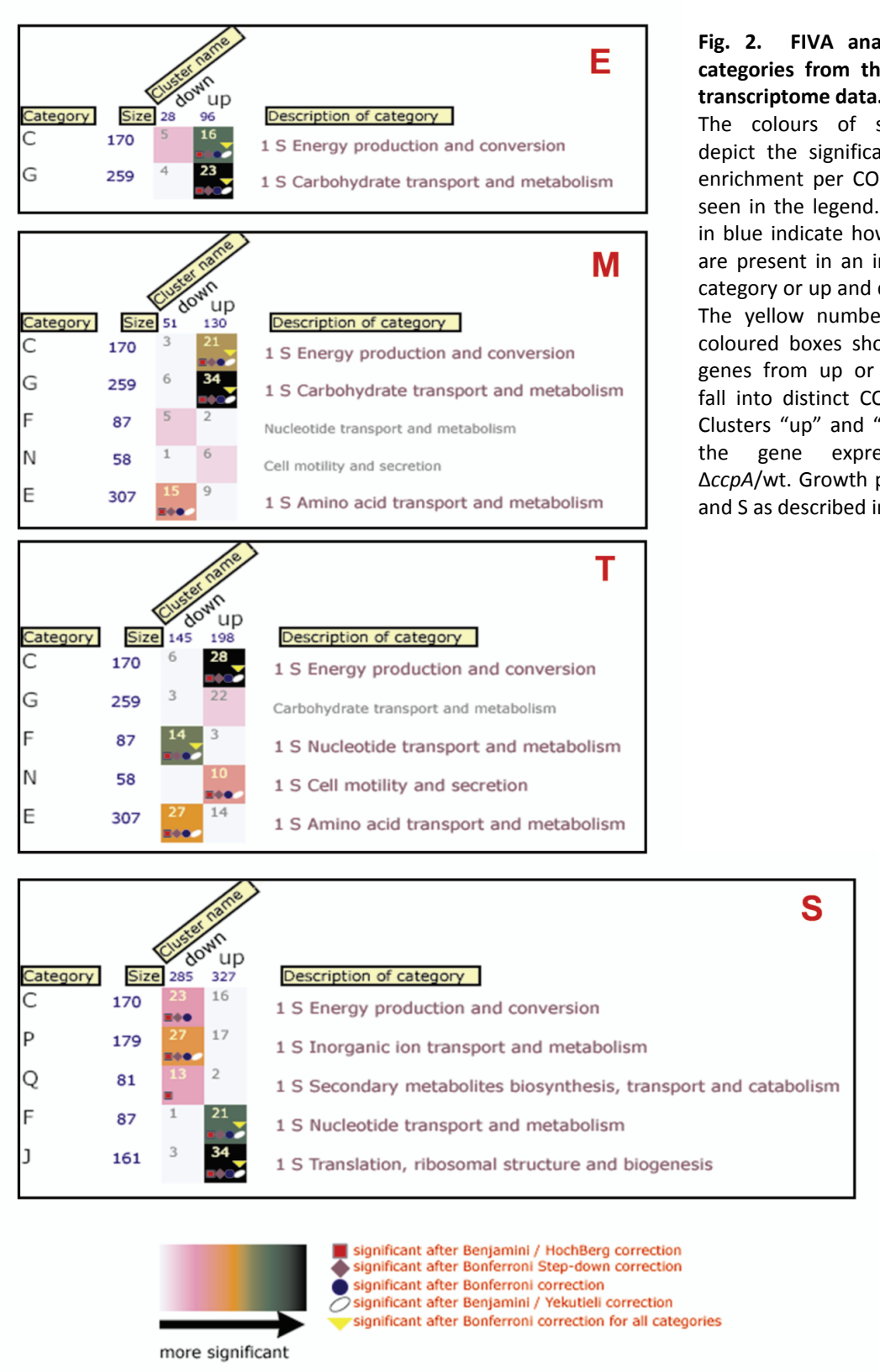


Fig. 2. FIVA analysis of COG categories from the time series transcriptome data.
The colours of square boxes depict the significance of genes enrichment per COG category as seen in the legend. The numbers in blue indicate how many genes are present in an individual COG category or up and down clusters. The yellow numbers within the coloured boxes show how many genes from up or down cluster fall into distinct COG categories. Clusters "up" and "down" reflect the gene expression ratios Δ*ccpA*/wt. Growth phases E, M, T and S as described in Fig. 1.

state the same COG's stay affected but with enhanced effects, as more genes fall into each individual cluster. There is an apparent switch in scope of the *ccpA* mutation after the cells enter the stationary phase of growth. Transcriptional profiles of two categories, namely 'nucleotide transport and metabolism' as well as 'energy production and conversion',

undergo reversion. In addition three new categories become influenced; (i) 'translation and ribosomal structure and biogenesis' with a majority of genes showing up-regulation; (ii) 'secondary metabolites biosynthesis, transport and catabolism' as well as (iii) 'inorganic ion transport and metabolism' with a majority of genes falling into cluster 'down'. The time-course approach indicates the dynamics of the CcpA modulon revealing an increasing range of CcpA regulation throughout the whole life cycle.

Clustering of time series data (STEM)

To identify expression profiles during the investigated time points the $\log_2$ transformed ratios from table S1 (supplementary material) were subjected to clustering using STEM [60] which implements a unique method of clustering of short time series. The data was divided into 15 arbitrary clusters with the maximum unit change set to one (ratio change of 2 in $\log_2$ scale) in model profiles between the time points. The obtained profiles were sorted based on the number of genes assigned (Fig. 3). A gene table with assigned profiles is available in the supplementary material (Table S2). Profiles 12 and 13, the former containing the highest number of 167 members, represent genes which are first up-regulated in the *ccpA* mutant in the logarithmic phase and down-regulated in the stationary phase. Profiles 1 and 5-9 depict genes which respond to lack of CcpA either by elevated (7, 8, 9) or by decreased (1, 5, 6) expression levels at specific stages of growth.

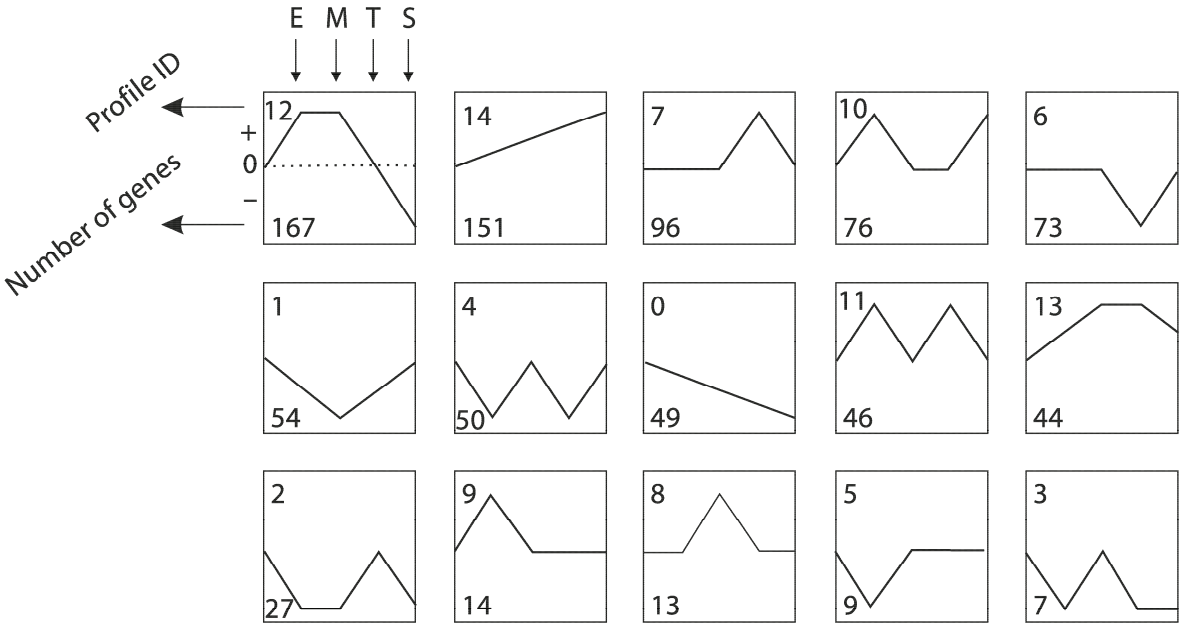


Fig. 3. STEM clustering profiles. 15 clusters contain the genes filtered on basis of the expression ratio change Δ*ccpA*/wt ≥ 1 (in $\log_2$ scale) between the time points. The dashed line indicates no change in expression between the strains. Growth phases E, M, T and S as described in Fig. 1.

Clusters 14 and 0 contain genes which respond to the absence of CcpA by constantly increasing or decreasing ratios, respectively. The remaining profiles (2, 3, 4, 10, 11) contain genes whose expression fluctuates in time. The general tendency of repression by CcpA, either direct or indirect, emerges from the distribution of genes, as the first four most densely populated clusters all have genes with higher expression levels in the *ccpA* mutant.

Regulation of the CcpA regulon members in time

We extracted known members of the CcpA regulon from DBTBS (**D**ata**B**ase of **T**ranscriptional Regulation in ***B**acillus **s**ubtilis*; [162]) and compared them to genes found to be regulated in our time series transcriptome data. The outcome of hierarchical clustering (complete linkage method) on this set of genes is presented in Fig. 4.

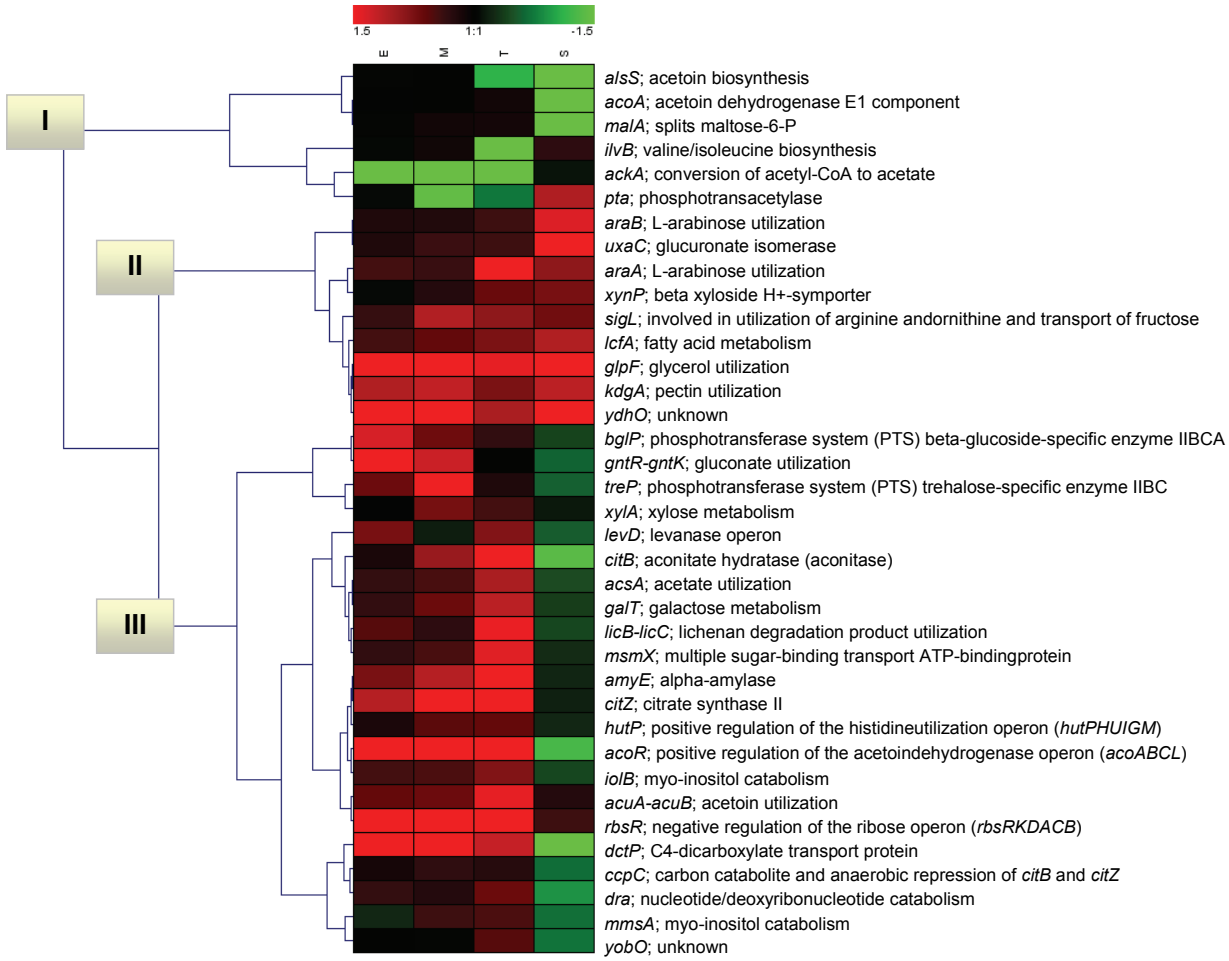


Fig. 4. Hierarchical clustering of regulation of the CcpA regulon members. Complete linkage clustering as an agglomeration rule was used in GENESIS programme. Red and green colour indicate higher or lower expression in the *ccpA* mutant strain, respectively. Growth phases E, M, T and S as described in Fig. 1. The three major trends emerging from the clustering are marked as I, II, III.

A few genes from this list were not regulated within our data set: *araE*, *bglS*, *citM*, *mmgA*, *yxjC*, *yxkJ*, *acuA*, *licB* and *gntR*. In case of the latter three, they were included for clustering since the other neighbouring operon member was significantly regulated (*acuB*, *licC* and *gntK*). In previous transcriptome studies [18,150,181,276] even more of the known CcpA regulon members, as listed in DBTBS, were missing. Three major trends can be distinguished within the clustering profiles including the down-regulated (I) and up-regulated (II) genes or up-regulated in exponential growth and down-regulated in stationary phase (III) (see Fig. 4). Genes involved in acetoin metabolism (*alsS*, *acoA*) seem to be CcpA-activated in later phases of growth, as well as another gene of overflow metabolism, *ackA*, which showed strong activation throughout the whole exponential phase. The *ilvB* gene, the first member of *ilv-leu* operon was specifically CcpA-activated at the transition state. On the other hand, *pta* was first activated in mid- and late exponential phase and subsequently repressed in the stationary phase. Genes from cluster III demonstrated an opposite expression pattern than the *pta* expression pattern. A few genes were repressed during all four growth phases, including, amongst others, *glpF*, *kdgA*, *rbsR* and *ydhO*.

Comparison of CcpA transcriptome data

Since four reports are already published describing the mid-exponential transcriptome of CcpA [18,150,181,276], we compared these data to ours (Table S1, supplementary material).
The comparison of five transcriptome data sets allowed us to identify a group of 53 genes (Table S3, supplementary material) common in at least three of the studies indicating that these genes likely belong to the CcpA regulon, although they are not present yet in the DBTBS database. 23 of these genes have a predicted *cre* site which makes them plausible candidates as direct CcpA targets. In case of *sucC, sucD, rocA* and *rocF* their protein products were also identified to be CcpA-dependent by a proteomics approach [276]. The comparison also provided an extraction of more than 200 genes not found in other whole-transcriptome studies [18,150,181,276] (Table S4, supplementary material).
Computational searches performed by Miwa et al. [173] resulted in the identification of 126 putative *cre* sites present in the *B. subtilis* 168 genome. Because different predictions for *cre* sites have been performed on the genome sequence of *B. subtilis* we decided to use a weight matrix (see supplementary material) based on the known *cre* elements of from the DBTBS database. With our search we detected 311 putative *cre* like sequences (Table S1).

Differentially regulated genes lacking a *cre* site

Analysis of the data in Table S1 revealed the presence of many genes which are higher or lower expressed in the *ccpA* mutant but lack the presence of a putative *cre* site. Expression of the whole *alb-sbo* operon coding for proteins required for the synthesis and maturation of subtilosin [287,288], the *dhb* operon (dihydroxybenzoate siderophore biosynthesis) [212] and *tdh* operon are strongly elevated in the transition phase. Also for *yerA*, the first gene of the

yecAyerABC operon, no predicted *cre* was found although it was found to be repressed at all time points in our studies as well as in two other transcriptome analyses [150,276]. On the other hand the expression of *manAyjdF* and *yqiX, mmgE* was decreased throughout the exponential phase and for some of these genes also in the stationary phase. These are just a few examples representing a large group of genes showing differential expression in the *ccpA* mutant but not possessing an identifiable *cre* sequence.

Comparison of glucose utilisation between the wild-type and *ccpA* mutant strain

It has been suggested that glucose uptake is slower in *ccpA* mutants [152,225,244] which could influence the glucose dependent regulation by CcpA especially in later stages of growth. Analysis of the concentration of glucose present in the medium showed that it was fully consumed by the wild-type strain at stationary growth, while it was still present in the culture of the *ccpA* mutant. With this knowledge in mind, we expected mainly CcpA-independent effects during stationary phase. In the stationary phase genes coding for proteins involved in the metabolism of nucleic acids and ribosomal proteins were strongly up-regulated in the *ccpA* mutant (Table S5), showing that most of these genes should be CcpA-independently regulated. However, some of these genes do contain a predicted *cre* sequence (for example *ggt, lcfA, sigY, ycdI, yknW, ysbA, yvdG, yxeB*) rendering them possible candidates of CcpA-dependent regulation independent of glucose presence. In these cases CcpA may interact with an unknown cofactor which is present intracellularly in the stationary phase or CcpA may bind to DNA without any associated cofactors.

Identification of new putative CcpA activated genes

The CcpA target *ackA* is one of the few genes that are activated throughout exponential growth but not in the stationary phase. Activation of *ackA* expression during growth in the presence of glucose is known to be dependent on the *cre* CcpA binding site (centred at -57 relative to the TS site) and an upstream activating region (UAR). Downstream of *ackA,* the gene for the molybdopterin precursor biosynthesis protein B (*moaB*) is located. *moaB* has an identical expression pattern as *ackA* suggesting it is in an operon with *ackA*.

A known target of CcpA, *pta*, is activated by CcpA during mid-exponential growth and in the transition phase, while repression was observed for the stationary phase. Its *cre* site is centred at -57-bp upstream of the TS. *yhbI* codes for a transcriptional regulator of the MarR family and it is most likely the first gene of the *yhbIJyhcABCDEFGHI* operon. *yhbI* is activated during exponential growth, while the downstream located genes *yhbJyhcAB* are also activated in the transition phase of growth. Moreno et al also observed activation of *yhbI* at mid-exponential growth in the presence of glucose [181]. A putative *cre* site was predicted around 57-bp upstream of a putative TS site. *manR,* which codes for the putative transcriptional activator of the mannose operon, is activated during exponential phase, while the putative *manPA-yjdF* operon is regulated not only at these time points but also during

stationary phase. Comparison of the promoter sequence of *manR* with those of the other CcpA activated genes we detected a putative *cre* site around 57-bp upstream of the putative TS. A second putative *cre* could be found which overlaps the TS site.

The gene for the proline transporter (*opuE*) is the only gene that was lower expressed in *B. subtilis* 168 compared to its *ccpA* mutant throughout growth. We detected a putative *cre* 104-bp upstream of the TS. The genes for the glycine betaine transport system components (*opuAABC*) [121] are activated at the first three time points. Two different TS sites were determined for this operon. A putative *cre* was found 143-bp upstream of the first TS site and another one 181-bp upstream of the second TS.

Identification of new putative CcpA repressed genes

Among the CcpA targets that were found to be repressed at all time points investigated were the known targets *glpFK*, *glpTQ*, *ycsA*, *yqgY*, *rocG*, the *ydh* operon and *sdhCAB*. Although Miwa et al. identified a putative *cre* located in the ORF of *sdhC* [173], we detected a putative *cre* site that overlaps the -35 sequence of the *sdhCAB* promoter. Interestingly *ysbA*, *yxjO* and the operons *yknWYZ* and *ptsGHI* were found to be repressed at the stationary phase only. All other were lower expressed in the wild-type at one of the exponential growth phases. For more than 50% of the cases we detected a yet unidentified *cre* site. The location of *cre* sites of all the exponential phase targets varied in overlapping either the -35, -10, TS or sequences upstream of the TS.

The *ydhMNOPQRST* and *kdgRKAT* operons each contain two *cre* sites. The first one is in both cases located between the TS and the start codon. The second one was found in *ydhT* and *kdgA*, respectively. Novel identified targets that are directly repressed by the presence of CcpA are the *ydhMNOPQRST* operon for which a putative *cre* site was detected present upstream of *ydhM* 7-bp downstream of TS and *ycgN* for which a *cre* site was identified within the ORF.

gene/operon	*cre position*	***cre* sequence**
sdhCAB	-40	ATGTACGCGTTTTCTT
glpFK	-27	TTGACACCGCTTTCAT
kdgA	ORF	ATGGAAGCGCTGACAT
glpTQ	4	AAGAAAGCGCTATCAT
ydh	7	ATGTAAGCGTTTTAAT
ydhT	ORF	TTGGAAGCGGTATCAT
phoPR	22	ATGAAAGCGCTATCAT
rocG	45	TTTAAAGCGCTTACAT
bglP	-36	ATGAAAGCGTTGACAT

Fig. 5. Alignment of the *cre* sites of genes that are CcpA regulated throughout growth. *cre* position reflects the distance from the centre of the *cre* to the transcription start site.

Location and composition of cre sites

Comparison of the promoter sequences of the CcpA-activated gene clusters showed that *yhbIJyhcAB* share a number of nucleotides suggesting a possible *cre* site and a UAR similar to *ackA* and *pta* [176] (Fig. 6). The distance from the centre of the *cre* towards the TS site, -57 bp, is identical in the promoter regions of *ackA* and *pta,* while for *yhbI* it could be located at -57 or 74-bp upstream of the putative TS. In the promoter region of *manR* two *cre* elements were identified of which one is located 57-bp upstream of the putative TS, determined on basis of the canonical -10 and hexameric -35 sequences. For the other activated genes/operons no such UAR was found.

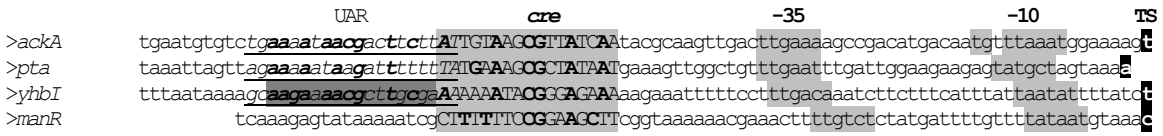


Fig. 6. Alignment of the promoter regions of the CcpA-activated genes. The upstream activating region (UAR) is underlined and in italics print. The *cre* sequence is in capitals and shaded in grey. The -10 and -35 sites are also marked in grey for each promoter. Transcriptional start sites are depicted by a white letter against the black background.

The *cre* sites of the CcpA-repressed targets were found to partly overlap the -35, -10, TS site or are located in the ORF (Table 1). Notably, genes repressed throughout growth (Fig. 5) share a conserved *cre* sequence. *bglP* is an exception as it has a very similar *cre* consensus, although it is repressed only in the early exponential phase.

Table 1 (next page). Genes and operons putatively regulated by CcpA by binding to a (predicted) *cre* sequence (in grey when predicted in this study). Genes located within one operon are listed from the promoter onwards. Gene expression ratios Δ*ccpA*/wt of activated (minus values) and repressed (plus values) genes are in bold when they meet the cut off of 2 and are indicated by 1 when the data did not meet the statistical criteria as described in Experimental procedures. The distance/location of the centre of the *cre* sites to the (in grey when predicted) transcription start is listed (Pos. *cre*1 or 2). Promoter characteristics such as the -35 and -10 sequences (both in capitals and italics), transcription start sites (in capitals, bold and underlined) and start codon (capitals and bold) are indicated in the (putative) *cre* sequences. Genes for which to promoter sequence could be predicted are indicated in light-grey at the end of the table. Growth phases E, M, T and S as described in Fig. 1.

Genes	E	M	T	S	pos. *cre* 1	pos. *cre* 2	Seq. *cre* 1	Seq *cre* 2
Activated								
opuAA	**-2.1**	**-2.4**	-1.8	1	**-143**	**-181**	aataaaagcgttttcaa	
opuAB	**-2.6**	**-3.7**	-1.9	1				
opuAC	**-2.1**	**-2.6**	1	1				
opuE	**-2.2**	**-2.4**	-1.5	-1.9	**-104**		tttgaaagcgttttatt	
ilvB	1	1	**-3**	1	**-88**		atgaaagcgtataca	
ilvH	1	1	**-4.9**	1				
ilvC	1	1	**-8.3**	1				
leuA	1	1	**-2.8**	1				
leuB	1	1	**-4.0**	1				
leuC	1	1	**-4.2**	1				
leuD	1	1	**-2.7**	1				
yhbI	**-2.6**	**-3.6**	1	1	**-74**		caagaaaacgcttgcgaa	
yhbJ	-2.0	**-2.7**	**-3.4**	1				
yhcA	**-2.4**	**-2.6**	**-2.3**	1				
yhcB	**-2.0**	**-2.8**	-1.9	1				
ackA	**-3.3**	**-2.9**	**-3.4**	1	-117	**-57**	attgtaagcgttcatcaat	attgtaagcgttatcaat
moaB	-1.9	-1.8	**-3.3**	1				
manR	**1**	**-2.4**	1	1	**-57**	**1**	ctttttttccggaagcttcg	aatgtaaac**G**gtttctta
manP	**-2.1**	**-3.8**	1	-1.9				
manA	**-6.8**	**-9.9**	1	-2.4				
yjdF	1	-1.6	1	-1.5				
pta	1	**-2.7**	-1.7	**2.1**	**-55**		tatgaaagcgctataatg	aa**A**gaaagcgtttttgta
Repressed								
ywcJ	1	1	1.9	1	**-143**		taagaatacgctttcatc	
citB	1	1.9	**4.9**	**-2.6**	**-106**		tgagagagcgtaaacagt	
levD	1	1	1.7	1	**-44**		aatgaaaacgcttaacaca	
levF	1	1	1.5	-1.78				
levG	1	1	2.0	**-2.2**				
sacC	1	1	1.5	1				
sdhC	**2.5**	**3.4**	1.7	**2.4**	**-40**	ORF	tatgtacgcgttttc*TTG*	TGTAACCGGTATCA
sdhA	1.6	**2.8**	1.6	1.8				
sdhB	1	**2.1**	1.8	1.9				
yqgY	1.6	1.7	**2.4**	1.5	**-37**		aatgaaaatg*TTTACA*tt	

Genes	E	M	T	S	pos. *cre* 1	pos. *cre* 2	Seq. *cre* 1	Seq *cre* 2
bglP	**2.5**	1.6	1	1	**-36**		aatgaaagcg*TTGACA*tc	
yxiE	**1.9**	1.9	1.9	1				
ysbA	1	1	1	**7.9**	**-36**		cttgtaagcgcttta*TAA*	
glpF	**3.3**	**3.9**	**2.7**	**4.3**	**-27**	ORF	a*TTGACA*ccgctttcatg	
glpK	**2.6**	**3.0**	**2.7**	**2.4**				
acoR	**3.7**	**4.8**	**5.2**	**-2.4**	**-27**		g*TTGAAA*gcgctttattt	
sspH	**3.7**	**3.2**	1	-1.9				
msmX	1	1	**2.6**	1	**-15**		taagaaagcgtt*TACAAT*	
sacP	**2.0**	**2.2**	1	1	**-2**		*GAT*gaaagc**G**tattctta	
sacA	**8.7**	**3.9**	1	1				
ywdA	**3.7**	1.5	1	1				
ycsA	1.7	1.6	**2.4**	-1.6	**-1**		atagaaag**C**gcttacgat	
sucC	**2.3**	**3.0**	**2.6**	1	**-1**		*AAT*gaaagc**G**cagtctat	
sucD	**2.6**	**2.0**	**4.4**	1				
yvfK	1	1.6	1	**2**	**3**		tctttaa**A**gcgctttcatt	
yvfM	**2.4**	**4.2**	1	1				
lacA	**3**	**3.7**	1	1				
glpT	**2.9**	**3.7**	**4.4**	**2.9**	**4**		taaga**A**agcgctatcatg	
glpQ	**2.6**	**4.4**	**5.9**	**5.5**				
amyE	1.7	**2.2**	**3.0**	1	**4**		aatgt**A**agcgttaacaaa	
yvfH	1	1	**2.9**	1	**4**		ataga**A**accgcttacat	
ytkA	1	1	**5.0**	1	**5**		gctg**T**aagcgtttgctac	
pmi	1.7	1	**2.1**	1	**5**		atag**G**aagcgttttctt	
ysfC	1	1	1.8	-2.0	**5**		attg**A**aagcgtttttga	
rbsR	**5.9**	**3.8**	**11.3**	1	**6**		tat**G**taaacggttacata	
rbsA	**3.9**	**2.7**	**3.6**	1				
rbsB	**5.5**	**3.3**	**6.7**	1				
rbsC	**5.4**	**3.7**	**5.9**	1				
rbsD	**4.5**	**3.0**	**7.4**	1				
rbsK	**2.4**	**2.0**	**3.4**	1				
ydhM	1	1.7	1	**4.6**	**7**		aa**T**gtaagcgttttaata	
ydhN	1	1.8	1	**-2.4**				
ydhO	**3.3**	**3.3**	1	**6.1**				

Genes	E	M	T	S	pos. *cre* 1	pos. *cre* 2	Seq. *cre* 1	Seq *cre* 2
ydhP	**2.5**	**3.3**	1	**8.1**				
ydhQ	**2.5**	**2.0**	1.8	**7.9**				
ydhR	1	**2.1**	1	**5.8**				
ydhS	1	**3.0**	1	**6.2**				
ydhT	1	1	1.8	**2.8**	**ORF**		attgaagcgcatactg	
yfiG	1	1	1	-1.5	**6**		ttagaaagcggttacaa	
yfiH	1	**2.2**	1	1				
ndk	**2.3**	1.8	**4.4**	1	**15**		tgtgaaagcctatacata	
kdgR	1	1	1	**3**	**17**		tttgaaatcggtttcaa	
kdgK	1	1	1	**2.7**				
kdgA	1	**2.3**	1	**2.2**	**ORF**		tggaagcgctgaca	
kdgT	1	1	**2.8**	**2.4**				
yxjO	1	1	1	**2.1**	**21**		catgtaaacgtaacaat	
odhA	**3.8**	**2.1**	1.9	1	**22**		gttggaagcgtttttatta	
odhB	**3.9**	**2.5**	1.9	1				
phoP	1	1.7	**2.5**	1.6	**22**		aatgaaagcgctatcata	
phoR	1	1	**2.2**	1				
yvdG	1	1	1.6	**-3.1**	**30**		gttgtaaccgctttctat	
cstA	1	1	1.8	1	**31**		a**ATG**aatgcggttacaat	
rocG	**3.4**	**5.6**	**10.3**	**3.2**	**45**		ttttaaagcgcttacatt	
yraO	1	**2.8**	**3.1**	1	**44**		tttgaaagcgcattcttg	
araA	1	1	**2.9**	1.8	**69**		tttgaaagcgttttattt	
araB	1	1	1	**2.6**	**ORF**		tgaaaacgattaca	
araL	1	1	1	1.7				
araP	1	1	1.6	1				
abfA	1	1.7	1	1				
citZ	**2.2**	**3.7**	**3.7**	1	**86**		aatgtaagcattttcttt	
mdh	1	1.5	1	1				
qcrA	**2.1**	**2.1**	**2.9**	1				
qcrB	1.8	1.8	**2.6**	1	**ORF**		tttgtaaccgttatccaa	
qcrC	**2.8**	**3.2**	**3.2**	1				
treP	1	**4.1**	1	1	**ORF**		tgtgaaaacgcttgcaga	
treA	**6.1**	**10.5**	1	**-2.4**				
treR	1	1.8	1	-1.6				

Genes	E	M	T	S	pos. *cre* 1	pos. *cre* 2	Seq. *cre* 1	Seq *cre* 2
xylA	1	1.7	1	1	**ORF**		tttggaagcgcaaacaaa	
gntK	**8.2**	**2.4**	1	1	**ORF**		tgaaagcggtacca	
gntP	**2.4**	1	**-3.5**	1				
gntZ	**2.3**	1	1	1				
rocD	**5.3**	**3.7**	**6.7**	1	**ORF**		gaggaattcgctgtcttt	
rocE	1	1.6	**3.7**	1	**ORF**		atcggaaccggtttctt	
rocF	**2.5**	1.9	**5.0**	1				
yneN	1	1.9	**3.7**	**-2.0**	**ORF**		ctttgaaaacgctatcagg	
acsA	1	1	**2.1**	1	**ORF**		cttgaaagcgttaccagc	
ydhT	1	1	1.8	**2.8**	**ORF**		aattgaagcgcatactgt	
yknW	1	1	1	**3.7**	**ORF**		aatggaaacgaatgtaga	
yknY	1	1	1	**3.8**				
yknZ	1	1	1	**4.0**				
spoVG	**2.4**	1.6	**3.9**	**-2.6**	**ORF**		catgagcgattgcatc	
sunA	**3.1**	**4.1**	1	1				
sunT	1	**6.1**	**2.1**	1	**ORF**			
bdbA	1	**3.7**	**2**	1				
yolJ	**3.4**	**4.8**	1.9	1				
bdbB	1	**6.5**	**2.5**	1				
ypiB	1	1	**2.4**	1				
ypiF	1	**2.6**	**4.3**	1	**ORF**		attgataccgctgtcata	
ycgM	1	1	1	**-9.7**				
ycgN	**6.5**	**6.0**	**8.7**	**-4.2**	**ORF**		acgtggaagcgtttaaaa	
ycgO	1	1	1	-1.8				
yxlH	1	**2.1**	1	1	**intergenic**		attgaaacgctttcaat	
yerD	1	1	**3.7**	1	**intergenic**		ctgtggaaatggttacata	
pckA	1	1	1.9	**-5.3**	**intergenic**		agagaaaccggttgctgg	
ptsG	1	1	**-2.2**	**2.7**	**intergenic**	ORF	aatgtaaacggttaaact	TGGAAGCGAATCCA
ptsH	1	1	1	**2.4**				
ptsI	1	1	1	1.7				
dctP	**4.9**	**6.9**	1	**-11**	**intergenic**		tatgaaaacgctatcatt	

Discussion

We compared the transcriptomes of wild-type and a *ccpA* mutant strain of *B. subtilis 168* at four different time points of growth in rich medium supplemented with glucose. The lack of the CcpA protein is sensed by the cells from early stages of growth as the expression of genes involved in carbohydrate metabolism was clearly influenced in the mutant. The effect developed further in time involving other functional categories of proteins when cells entered the later stages of exponential growth as shown by FIVA analysis. These effects are caused by direct or indirect CcpA action and that is why a distinction should be made between the CcpA regulon and the CcpA-modulon, respectively. The CcpA-modulon would encompass the genes whose regulation results, for example, from the altered cell physiology due to slower glucose uptake caused by lack of functional CcpA. Indeed, many genes with no apparent *cre* site were differentially expressed in our, as well as in other transcriptome studies [18,150,181,276] which make them likely candidates for the CcpA-modulon. However, another yet unknown mode of regulation by CcpA could be present which would allow classifying at least some of these genes as direct CcpA targets. The *tdh* and *yerA* genes are just two interesting examples. They have no predicted *cre* sites but show a clear repression throughout time in our experiments.

In addition, the *cre* sequence itself is still not precisely defined yet, since the search performed with the WTGNAANCGNWNNCW consensus [173] retrieved only 126 genes with possible *cre's* but still novel functional or putative *cre's* were reported [110,171] and an altered *cre* consensus WTGAAARCGYTTWW (W=A or T; N=any base; R=A or G; Y=C or T) was proposed [171]. Taking this into account and the fact that DNA binding properties of CcpA may differ at different stages of growth, we performed our own *cre* search within the whole *B. subtilis* genome using known *cre* sites from DBTBS and applying a non-stringent, 0.85 score cut-off. This search resulted in more than 300 hits. *In vitro* studies [79] have shown that CcpA binding to *cre* sites is dependent on the acidity of the buffer used. It is possible that during growth on a rich glucose medium internal pH changes in time, which could influence recognition and CcpA-binding to its target sequence. Moreover, the availability and concentrations of various cofactors change in time, which would also affect CcpA-DNA binding or the CcpA interactions with the transcription machinery. The latter mechanism has been already suggested [126,127] and a NADP molecule appears to be one of the putative coeffectors in this process [126,127]. Under our experimental conditions the CcpA protein in the wild-type strain was present at all stages of growth (Western blot analysis, data not shown), even in the stationary phase when glucose was already absent from the medium. This demonstrates that CcpA is constitutively expressed and corresponds well with previously published data [174]. Most likely the temporal regulation of different groups of genes by CcpA results from the difference in quantitative and qualitative availability of cofactors, the presence of which would orchestrate alternative modes of CcpA action. In the

exponential phase of growth the derivatives of glucose metabolism probably modulate CcpA interaction and when glucose is depleted other coeffectors could take over the role in CcpA-dependent gene regulation. The interplay of two distinct coeffectors, G6P and HPr-Ser-P, has been already shown for the *gnt* operon [172] and the *xyl* operon [79]. In the latter case the acidity played a role in selective determination of the CcpA interaction with cofactors. Besides, both operons contain multiple *cre* sites present in the promoter regions and structural genes. We have identified several genes and operons also containing more than one *cre* site, whose expression was altered when comparing the wild-type and the *ccpA* mutant strain. These included*: ptsG, sdhC, glpF, kdg, rbs* and the *ydh* operon. The *ydhMNO* genes are involved in the transport of oligo-β-mannosides [207].

The *ilvBHC-leuABCD* has been also shown to be activated by CcpA [225,244]. This operon is repressed by CodY and TnrA, but during transition phase the operon is activated in the presence CcpA in the wild-type strain. This is likely due to the shortage of amino acids or GTP, which results in a loss of the repression by CodY. This shows another aspect of temporal regulation by CcpA, namely interplay of distinct regulators of gene expression whose presence differs in the course of time and bacterial growth and depends on the availability of energy sources.

Another interesting observation emerging from our time series data concerns the strong activation of the *yvd* gene cluster and the *malA-yfiA-malP* operon exclusively in the stationary phase. This coincides with the lack of repression of *amyE,* the expression of which was repressed in the exponential phase. The α-amylase, coded by the *amyE* gene, hydrolyzes extracellular polysaccharides to smaller sugars, such as dextrins, maltose and glucose which can be transported through the cell membrane. The *mal* operon and the *yvdE-pgcM* cluster (*yvdGHI=mdxEFG*) are responsible for the maltose and maltodextrin utilisation, respectively [218,272]. Glucose repression of these genes has been shown to be mediated by CcpA and *cre* [272]. Interestingly, the *yvdE-pgcM* gene cluster contains a *cre* in the *yvdG* (*mdxE*) intergenic region and shows the same expression pattern in time as the *mal* operon, suggesting that CcpA is also involved in the regulation of the maltodextrin utilisation.

With our approach we have identified three new gene clusters, namely *yhb*, *man* and *opu,* and one gene, *manR,* that are activated by CcpA. Alignment of the *cre* site and its UAR of the well-studied CcpA target *ackA* showed a certain degree of homology with the promoter region of *yhbI* (Fig. 6), suggesting a similar type of regulation. No homology with the UAR was obtained with the promoter region of *manR,* which has a putative *cre* located at the same location as *ackA*. Assuming that the *cre* site of *yhbI* is located at -57 bp from the putative TS site, all these activated genes would follow the model proposed by Kim et al. [127]. These authors have shown that the CcpA binding site has to be located at the same site of the helix as the TS site. Except for *opuE,* activation of expression of all the CcpA activated genes was lost in the stationary growth suggesting that activation might be strictly glucose dependent.

In this study, we show dynamic regulation by CcpA in time. Analysis of the data has identified potential new members of the CcpA regulon and hints at potential differential

binding specificity that may depend on the relative presence or absence of perhaps yet unknown cofactors. We show that some transcriptional effects might be attributed to a difference in glucose consumption, though also glucose-independent regulation by CcpA may occur.

Supplementary material

Supplementary tables and graphs mentioned in this manuscript as well as the slide images and raw data are available from http://molgen.biol.rug.nl/publication/ccpa_bsub_data/

Acknowledgements

The authors thank Evert-Jan Blom for providing the possibility of using the FIVA software before publication and Sacha van Hijum for his support in the microarray data handling process and the preparation of the website with supplementary material. We are also grateful to Aldert Zomer for the skilful support with MotifLocator, Wiep Klaas Smits for helpful discussions and Siger Holssappel for technical assistance in determining the glucose concentrations.

This work was supported by IOP genomics grant IGE01018 from the 'Innovatiegerichte Onderzoeksprogramma's' of the Dutch Ministry of Economic Affairs.

Chapter 6

Regulon definition of CcpA mutants with altered activities in *Bacillus subtilis*

Andrzej Lulko, Siger Holsappel, Mareen Sprehe, Wolfgang Hillen, Girbe Buist and Oscar Kuipers

ABSTRACT

Carbon catabolite regulation is an important and well-characterized process, mainly under control of CcpA, a pleiotropic regulator in low-G+C Gram-positive bacteria. In this study, transcriptome analysis was applied to investigate the effects of point mutations in CcpA on gene activation or repression at mid-exponential and stationary growth phase of *Bacillus subtilis*. The target amino acid residues were localized either in the DNA-binding domain (mutant residue R54W) or in the HPr(CrhP)-Ser-P binding domain (mutant residues Y90P, V301W and T307W).

The transcriptome profiling revealed that the R54W, Y90P and V301W mutants have a severely reduced, but not completely absent, capability to perform carbon catabolite regulation, whereas the T307W strain has a less impaired activity, especially in relation to the known CcpA regulon members. Some deviations to this general trend were observed, although no clear explanation based on the analysis of the location and the sequence of catabolite responsive elements could be deduced. Based on the comparisons between the individual point mutants, the T307W mutant not only has a substantially different overall transcriptome profile compared to the wild-type, but also shows hardly any overlap in the affected functional categories, when put side by side with the other three mutants. The other strains had altered CcpA activity on genes belonging to different functional categories, such as amino acid and carbohydrate transport and metabolism as well as energy production and conversion. Furthermore, the results obtained by the microarray approach were compared with the phenotypic behaviour of all mutated strains and the CcpA wild-type counterpart by measuring the rate of glucose consumption from the medium. The R54W, Y90P and V301W mutants did not utilize glucose in a preferential manner, while the T307W variant utilized glucose similarly to the wild-type.

The results presented here suggest that the particular mutation T307W has a modest influence on the functioning of CcpA, whereas the other three mutations, especially Y90P, severely affect activity of this transcriptional regulator. The transcriptional activities of the mutants were confirmed at the phenotypic level as shown by the glucose utilization experiments. Residue Y90 was shown to be of crucial importance for CcpA functioning, probably because when this residue is mutated CcpA loses the ability to bind HPr-Ser-P, while V301 shows this behaviour to a lesser extent.

Introduction

Carbon catabolite regulation (CCR) is a regulatory mechanism that allows cells to utilize the available carbon sources in a preferential and optimal order. Carbon catabolite control protein A (CcpA) is the master transcriptional regulator of the CCR pathway in Gram-positive bacteria [235]. This regulator directly or indirectly activates or represses expression of several hundreds of genes during the growth cycle of *Bacillus subtilis* as has been shown by transcriptome analysis using samples form different growth phase [154]. Pleiotropicity of *ccpA* knock-out mutants has been confirmed by various transcriptomic studies [18,150,181,276]. Independent of the availability of preferred carbohydrates in the growth medium, CcpA is constitutively expressed [174], which implies that this regulator needs additional factors to exert its pleiotropic function. Indeed, DNA-binding activity of CcpA is stimulated by complex formation with HPr-Ser-P [52],[75] or HPr-like protein Crh-Ser-P [77]. Also low molecular weight molecules, such as NADP, glucose-6-phosphate (G6P) and fructose-1,6-bisphosphate (FBP), can modulate either DNA-binding properties or the interaction with the transcriptional machinery of the CcpA-(HPr-Ser-P) complex [79,125,126,223]. Genes subject to CcpA-mediated CCR are in the vicinity of a specific binding site, known as a catabolite-responsive element (*cre*), which can be located either within ORFs or in intergenic sequences.

The crystal structures of the ternary complexes of (HPr-Ser-P)-CcpA-DNA and (Crh-Ser-P)-CcpA-DNA are available for *Bacillus megaterium* [219,220]. These structures revealed that CcpA is composed of an N-terminal DNA binding domain (residues 1-60) and a C-terminal domain that contributes to dimerization and corepressor (HPr or Crh) binding. The former domain binds to the DNA sequence through the helix-turn-helix (HTH) motif and the hinge helix, which also connects to the C-terminal domain. These two motifs make 32 phosphate contacts to 10 bp out of 16 bp of the *cre* site used, as shown by the crystal structure analysis. The most important contacts are formed with the central GC pair, G(3) and C(14) of the pseudopalindromic *cre* sequence. On the other hand, the crucial role of HPr-Ser-P binding is attributed to Tyr89 (corresponds to Tyr90 in *B. subtilis*) and the residues 294-309 (helix IX) of CcpA, with leading roles of Tyr295, Ala299, Val300 (Val301 in *B. subtilis*), Arg303, Leu304 and Lys307. Upon HPr binding N-terminal subdomains rotate, leading to the juxtaposition of the DNA-binding regions of the CcpA dimer that results in "hinge" helix formation [219]. This so-called allosteric switch mechanism in the presence of cognate DNA is a central event in HPr-Ser-P-dependent CCR.

Recently, several variants of CcpA containing single amino acid exchanges located in different regions of CcpA of *B. subtilis* were investigated with regard to their transcriptional activities on two CcpA-activated genes (*ackA*, *alsS*) and two repressed ones (*xynP, gntR*) [231]. These experiments revealed that the CcpA strains mutated in the HPr(CrhP)-Ser-P-binding regions were inactive in CCR of these four genes. On the other hand, the CcpA

variants impaired either at the amino acid residues that change the CcpA conformation upon corepressor binding or at the residues in the hinge helix, showed a variety of complex regulatory effects, which were dependent on the promoter as well as the type of the mutation. We performed transcriptome analyses of four strains each harbouring a different CcpA point mutantion at mid-exponential and stationary phase of growth, in rich glucose-containing medium. The aim of this study was to evaluate the effects of these CcpA point mutations, which are localized in either the hinge helix (R54W) or the HPr(CrhP)-Ser-P binding domain (residues Y90P, V301W and T307W), on the regulon definition of CcpA.

Materials and methods

Bacterial strains, media, growth conditions and preparation cells

Four CcpA point mutants used in this study were prepared via a two-step PCR mutagenesis method using the pWH1541 plasmid as the template as described previously in detail by Sprehe *et al.* [231].

Table 1: Strains and plasmids used in this study.

Strains	***Characteristics***	***Reference***
B. subtilis WH440	Δ*ccpA*; *trpC2*, *amyE*::(*xynP' -lacZ cat*)	[223]
B. subtilis ccpA::spc	*trpC2 ccpA::Tn917 spc*	[154]
Plasmids	**Characteristics and description**	
pWH1541	His_6-*ccpA* (wild-type reference strain)	[231]
pWH1541-R54W	Derivative of pWH1541 His_6-*ccpA* R54W	[231]
pWH1541-Y90P	Derivative of pWH1541 His_6-*ccpA* Y90P	[231]
pWH1541-V301W	Derivative of pWH1541 His_6-*ccpA* V301W	[231]
pWH1541-T307W	Derivative of pWH1541 His_6-*ccpA* T307W	[231]

The *ccpA* deletion mutant WH440 [223] was transformed with pWH1541 (containing his-tagged CcpA and referred to as wild-type here) or different versions of pW1541 [231] plasmids which contained the mutated proteins as indicated in Table 1. Cultures were grown in 250 ml Erlenmeyer flasks by inoculation with "synchronized" cells at OD_{600} 0.1 in the 25 ml TY medium supplemented with 1% glucose. Synchronized stocks for inoculation were prepared by diluting overnight (ON) cultures to OD_{600} 0.01 in fresh medium and growing the cells to the mid-exponential phase, and diluting again to OD_{600} 0.01. After three cycles of growing to mid-exponential phase synchronized stocks were frozen in glycerol at -80°C for future use. The reason for synchronization was to avoid variable lag-phases of the different CcpA mutants, which were observed when direct inoculation from ON cultures was

attempted. Cells for RNA isolation and subsequent microarray analysis were harvested by centrifugation (1 min, 12,000 rpm, 4°C) of mid-exponential phase as well as stationary phase (approximately 2 hours after transition point) cells corresponding to approximate OD_{600} values of 0.5 and 4.5, respectively. Cell pellets were immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation.

DNA microarray experiments

Preparation of total RNA, cDNA synthesis, fluorescent labeling, DNA microarray hybridization, washing and scanning as well as image analysis, data normalization and the statistical analysis of global gene expression were performed as described by Lulko *et al,* [154]. The ln-transformed ratios of the expression levels of *ccpA* mutant versus wild-type were subject to a t-test using Cyber-T [149].

To identify biological processes influenced by the CcpA mutations, transcriptome data analysis was performed with FIVA [20]. A combination of "threshold" and "iterative Group Analysis" methods within FIVA was implemented to carry out partitioning of the data into two groups (down-regulation and up-regulation corresponding to repression and activation, respectively) with the threshold borders set to 2 for both groups (see http://bioinformatics.biol.rug.nl/standalone/fiva/fivaTutorial.php#BB and [25] for details).

Expression diagrams were generated in Genesis by sorting on the mean expression values [236]. The same clustering software was used to perform principal component analysis (PCA) based on the mutant samples to evaluate general trends in the data as well as to explore correlations among the individual CcpA species.

Glucose utilisation

D-glucose depletion from medium during growth was determined at one hour intervals for the wild-type and the *ccpA* mutant strains by means of the D-Glucose enzymatic bioanalysis/food analysis kit (R-biopharm) as described before [154]. The experiments were performed in duplicate.

Results

Characterisation of growth and glucose consumption

The DNA microarrays were used to compare the transcriptional profiles of four *B. subtilis* CcpA point mutant strains (R54W, Y90P, V301W and T307W) with the corresponding wild-type strain *B. subtilis* WH440 (pWH1541) containing plasmid encoded CcpA. Samples for DNA-microarray analyses were collected at the mid-exponential (2 hours before the transition phase of growth) and at the stationary phase (2 hours after transition, when most glucose is depleted), which corresponded to an optical density of the cells of 0.5 and 4.5,

respectively. The growth characteristics and glucose consumption were examined in the rich TY medium supplemented with 1% glucose. The difference in growth rate between the strains during exponential growth was negligible (Fig. 1), however, the mutant cultures, especially V301W, reached lower OD_{600} values at the end of the stationary phase (after 7 hours of growth).

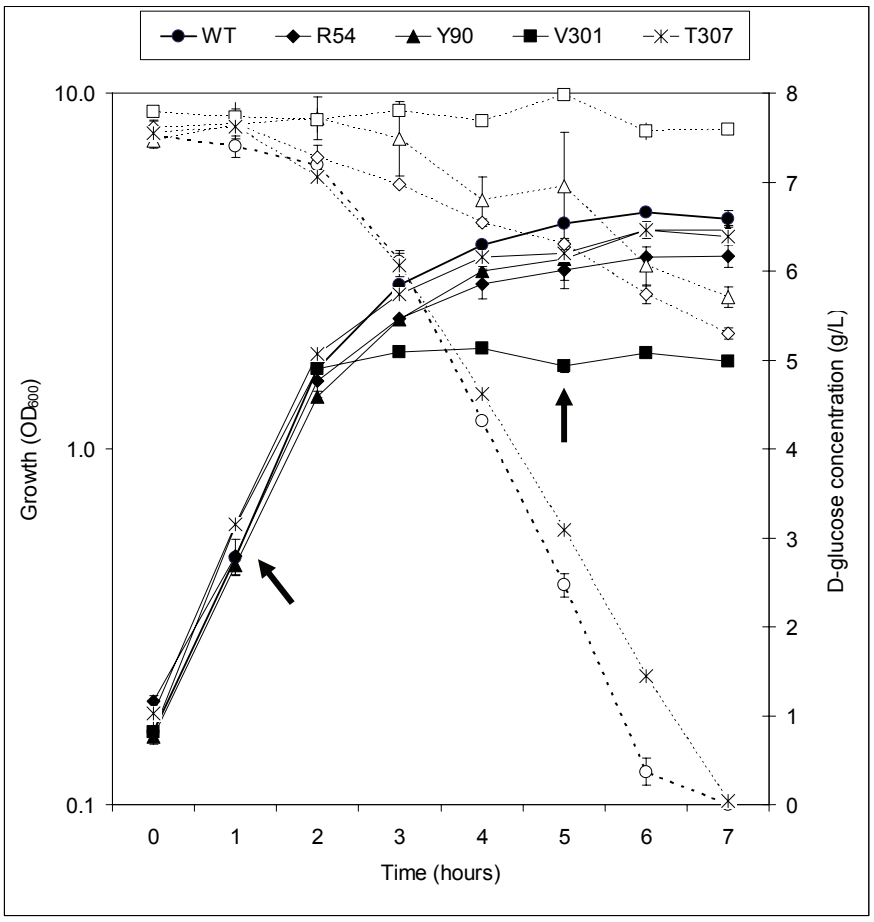


Figure 1. Growth characteristics (solid lines) and glucose utilisation (dashed lines) of *B. subtilis* wild-type (WT) strain (pWH1541) and the CcpA point mutant strains (see legend for details). The OD and glucose measurements were performed at hourly intervals. Arrows indicate sampling time for the transcriptome analyses.

The glucose level of the wild-type strain started to sharply decrease after 3 hours of growth, whereas in the case of the mutant strains this decrease was either not detectable (V301W) or very slow (R54W, Y90P). In contrast, the level of glucose of the T307W mutant began to drop from the third hour on, though it was slightly slower when compared to the wild-type strain. Although this difference in glucose concentration in the medium of individual mutants will not affect the transcriptome analysis in the exponential phase of growth, it is expected to have a significant effect on the regulation by CcpA in the stationary phase.

Functional categories of genes affected in expression in the individual mutants

The transcriptome data obtained for the mutant and WT strains at different time points are presented in the form of Lowess-normalized *ccpA*_mutant/wt expression ratios the

significance of which was evaluated with the Cyber-T statistical package [12]. Although the protein levels of CcpA in all tested mutant strains were comparable as shown previously by Sprehe *et al.* [231], ensuring that the obtained results can be truly attributed to the characteristics of the mutants and do not result from the differences in the CcpA expression levels among the strains, our transcriptome data revealed that in the case of Y90P and T307W the CcpA mRNA levels were decreased by a factor of 2.2 and 2.5 in these two strains, respectively.

When comparing a mutated CcpA strain against its wild-type counterpart, genes that show either activation or repression indicate that a given mutation renders the CcpA protein (partially) inactive. Oppositely, if there is lack of regulation difference (ratio ~1), it means that this point mutation has a harmless effect on the structure and function and that the mutated CcpA protein has an activity comparable to the wild-type protein with respect to this gene. Since ratios of the expression levels of a given *ccpA* mutant versus wild-type are presented here, thus a ratio above 1 (up-regulation) indicates repression of a particular gene as the expression level in a mutant was higher (also called derepression) than in the wild-type strain due to lack of the functional repressor. Oppositely, a ratio below 1 (down-regulation), presented here as negative values, indicates activation of a particular gene as the expression level in a mutant was lower due to lack of the functional activator.

To identify biological processes influenced by the CcpA mutations, the FIVA [20] software was used to screen the gene expression data for significantly overrepresented clusters of orthologous groups of proteins (COG) and gene ontology groups (GO), both of which contain functionally related genes, using the settings indicated in the Material and Methods. Three of the mutants (R54W, Y90P and V301W) and the full-knockout strain (time-resolved transcriptome analysis of which is presented in chapter 5) do not drastically differ in their functional profiles as the genes belonging to 'amino acid' and 'carbohydrate transport and metabolism' as well as 'energy production and conversion' COG's are significantly altered in their expression profiles (Table 2). The processes like 'branched-chain amino acid biosynthesis', 'carbohydrate transport' or 'TCA cycle' showed the same pattern of regulation in the full-knockout strain as in the R54W, Y90P and V301W mutants. Still, a subtle divergent behaviour of particular strains for the applied threshold limits was also detected, for example genes for 'acetoin biosynthesis' (*alsDS*) or 'amino acid transport' were only affected in the V301W and Y90P mutants, respectively. Up-regulation of 'ATP synthesis coupled proton transport' (GO_0015986) was observed in the case of the R54W and V301W strains and the Y90P mutant showed an exclusive effect on 'amino acid transport' and 'pyrimidine metabolism'. The V301W mutant did influence neither 'electron carrier activity' nor 'threonine biosynthesis', whereas the R54W and Y90P mutant did. Last but not least, the T307W strain is clearly unique as it shows little influence on CcpA-specific transcriptional activity different from the wild type strain. However, this mutant still showed differentially expressed genes (some examples are given in Table 5) when compared to the wild-type

strain (intact CcpA). Many of these genes do not belong to the CcpA regulon (data not shown) as defined in the DataBase of Transcriptional Regulation in *B. subtilis* (DBTBS) [162].

Table 2. Summary of the overrepresented clusters of COG and GO groups of the CcpA point mutant strains in the exponential phase. Empty boxes indicate no effect. Up-regulation (up, dark grey boxes) of a particular category indicates that the expression levels of genes belonging to this category were higher in the mutant (derepression as an effect of the inactive repressor) than in the wild-type strain. Down-regulation (down, light grey boxes) of a particular category indicates that the expression levels of genes belonging to this category were lower in the mutant (due to the inactive activator) than in the wild-type strain.

COG/GO category description	**COG** / GO	KO_M*	KO_T*	**R54W**	**Y90P**	**V301W**	**T307W**
Amino acid transport and metabolism	**E**		down	down	down	down	
Carbohydrate transport and metabolism	**G**	up	up	up	up	up	
Energy production and conversion	**C**	up	up	up	up	up	up[3)]
Nucleotide transport and metabolism	**F**				up	down	
Acetoin biosynthetic process	0045151		down			down	
Acetoin catabolic process	0045150		up				
Amino acid transport	0006865	down			down		
Amino acid biosynthetic process	0008652		down	down	down	down	
ATP synthesis coupled proton transport	0015986			up		up	
BCFA[2)] biosynthetic process	0009082		down	down	down	down	
Carbohydrate transport	0008643	up	up	up	up	up	
Electron carrier activity	0009055		up	up	up		
Glycerol metabolic process	0006071	up	up	up	up	up	
NAD biosynthetic process	0009435						down
Pyrimidine nucleotide metabolic process	0006220				up		
Sugar:hydrogen ion symporter activity	0005351	up	up	up	up	up	
TCA cycle	0006099	up	up	up	up	up	up[3)]
Threonine biosynthetic process	0009088		down	down	down		

[1)] SM – secondary metabolites ; [2)] BCFA – branched chain family amino acid; [3)] regulation was less pronounced when compared to the other strains;

* KO_M and KO_T refer to data of the *ccpA* knockout strain at the mid-exponential and the transition stage of growth, respectively ([154]**.**The sampling OD for the KO-M and KO-T samples was 0.8 and 3.4, respectively, while for the exponential samples of the single amino acid mutants of CcpA was 0.5.

Significantly affected functional categories of the T307W mutant included transport/binding proteins and lipoproteins, membrane bioenergetics as well as metabolism of carbohydrates, amino acids and related molecules (data not shown). Besides, 75 genes, which corresponds to approximately 35% of the total regulated genes in T307W, have no assigned function. This goes in line with all the other evidence presented within this manuscript and strongly suggests that with respect to the functional CcpA this strain is the most similar to a wild-type like phenotype.

The same type analysis applied on the data derived from the stationary phase of growth revealed a much more diverse distribution of the regulated GO's as well as a higher number of affected COG categories (data not shown). In addition to the categories disclosed already

in the analysis of the first time point of growth, new categories have been identified at the later stage of growth, such as down-regulation of 'cell division and chromosome partitioning' (R54W), 'cell motility and secretion' (V301W) or 'translation, ribosomal structure and biogenesis' (R54W). Similarly to the outcome for the exponential phase of growth, the profile of the T307W mutant was fairly distinct from the other strains (data not shown).

While comparing the full-knockout strain to the point mutants strains some discrepancies in the differently affected categories were noted. These resulted most likely from the fact that the microarray experiments with the full knockout strain were performed with lower concentrations of glucose, *ie*. 0.5% instead of 1%, and this was already absent from the medium at stationary phase in the previously published experiments ([154]), chapter 5), whereas in this study glucose was still present in the stationary phase of growth (Fig. 1).

Comparison of the mutant transcriptome data for the known CcpA regulon members

To determine the effect of the CcpA point mutations on the expression of the CcpA regulon the transcriptome data of the mutants were compared for the regulon members as defined in DBTBS [162] Expression diagrams of the mean expression values and principal component analysis (PCA) to evaluate general trends as well as to explore correlations among the individual CcpA species (both generated in Genesis [236]) allowed us to draw the following conclusions.

(i) For three of the mutants (R54W, Y90P and V301W) and the full-knockout strain the vast majority of the genes showed an apparent effect (*ie*. repression or activation, indicated as red and green bars in Fig. 2, respectively). Especially the Y90P mutation has a very strong CCR phenotype that is comparable with the deletion mutant. This implies that the point mutants have a phenotype similar to the full *ccpA* knockout, indicating that the mutation rendered CcpA (almost) inactive. Genes which showed little or no effect in one of the phases of growth included: *ccpC*, *mmgA*, *licB*, *levD*, *acoA*, *malA*, *araE* and *yxjC*. Since for these genes the regulation was also not present in the case of full-knockout, it can be assumed that under the conditions used here either CcpA does not act on these genes or the level of transcripts could not be properly measured by the DNA microarrays.

(ii) Repression is predominant as compared to activation, which is a well-known phenomenon based on the CcpA literature. This predominance is particularly clear in the mid-exponential phase and becomes less obvious in the stationary phase.

(iii) Additionally, the PCA comparisons revealed that this kind of agglomeration can not be clearly demonstrated for the stationary phase. The scattered PCA pattern can be probably attributed to different concentrations of glucose present in the media of the investigated strains in the stationary phase of growth.

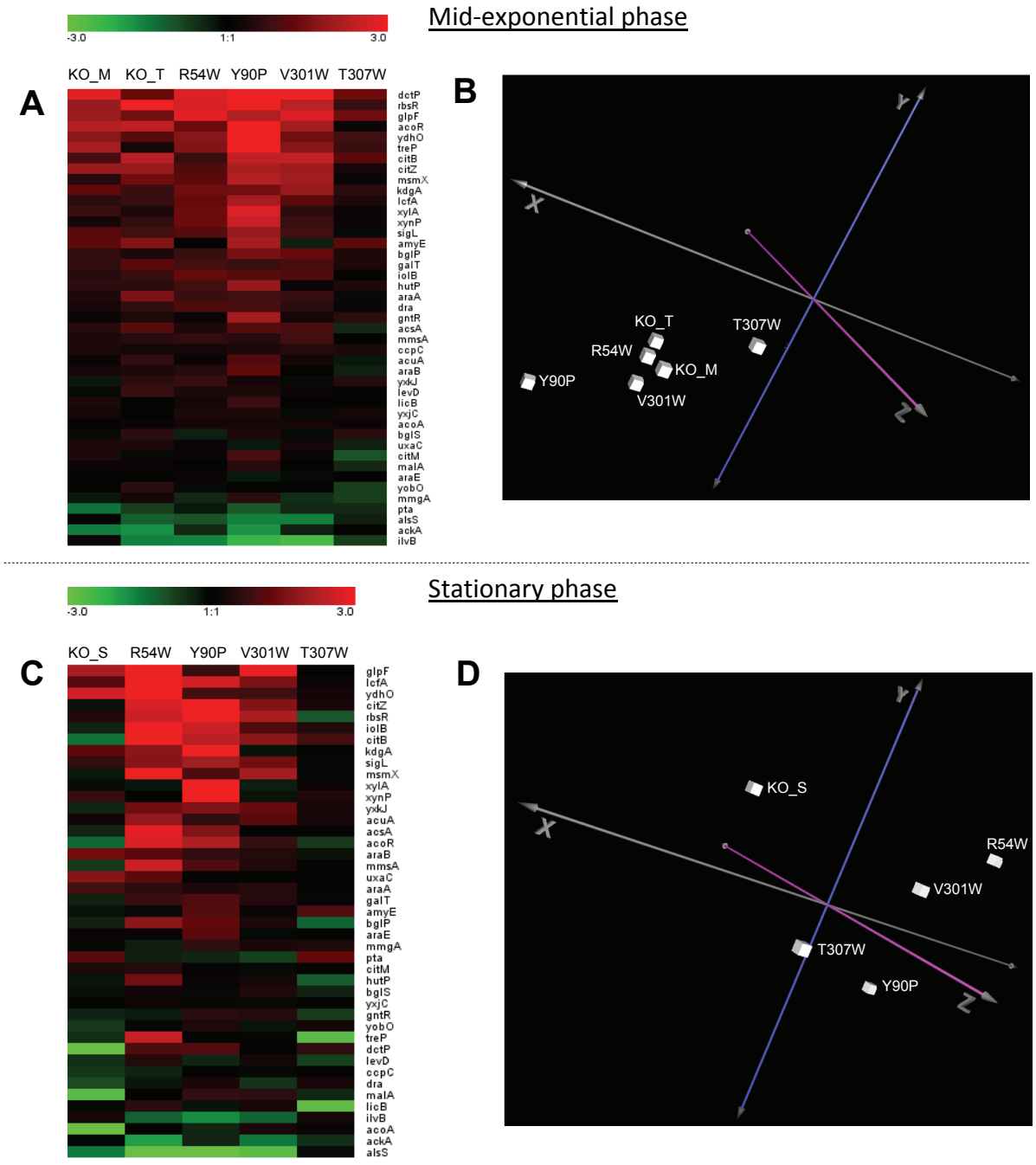


Figure 2. Regulation of the CcpA regulon members. Expression profiles (A, C) and the PCA graphs (B, D) were generated in the Genesis software on the log^2 transformed gene expression ratios (mutant vs wild-type). Red and green colour indicates higher or lower gene expression, corresponding to repression and activation in the individual *ccpA* mutant strain, respectively.

As demonstrated in Fig. 2D, the mutants are definitely more scattered when compared to the situation in mid-exponential phase. Concordantly, a hierarchical clustering approach revealed that the T307W mutant is the least similar to the full knock-out strain and that the R54W, Y90P and V301W strains co-cluster together (data not shown) which confirms the pattern obtained by the PCA approach. Still, as described below, several other genes not

listed in DBTBS showed differential expression in the T307W strain, indicating that, although limited, the mutation had impact on the expression of other genes (see also Table 5).

Table 3. Regulation of known CcpA-activated genes in the CcpA mutants. Numbers represent gene expression ratios (*ccpA* mutant/wt).

Gene	KO_ M*	KO_ T*	R54 W	Y90 P	V301 W	T307 W	KO_ S*	R54 W	Y90 P	V301 W	T307 W
	Mid-exponential phase						**Stationary phase**				
ackA	**-2.9**	**-3.4**	-1.4	**-3.5**	-1.3	1.0	-1.1	**-3.9**	-1.4	**-2.7**	-1.5
alsS	1.0	**-2.2**	**-2.0**	**-3.0**	**-3.0**	-1.3	**-2.8**	**-12.6**	**-15.2**	**-6.8**	-1.1
pta	**-2.7**	-1.7	-1.2	**-2.1**	-1.4	-1.4	**2.1**	-1.3	-1.4	**-1.7**	**2.4**
ilvB	1.1	**-3.0**	**-3.1**	**-5.1**	**-5.9**	**-1.7**	1.2	**-2.3**	**-3.3**	**-2.3**	1.2
opuAA	**-2.4**	-1.8	-1.8	**-4.1**	-1.1	1.6	1.1	-1.7	**-2.1**	-1.5	**-2.3**
opuE	**-2.4**	-1.5	-1.8	**-2.6**	**-2.4**	**-2.4**	-1.9	1.1	**-1.9**	1.1	-1.3
yhbI	**-3.3**	**-4.7**	-1.2	**-3.5**	-1.2	-1.4	1.5	**-2.2**	**-2.7**	**-2.7**	**-2.3**
manA	**-9.9**	**-2.2**	1.3	**-2.0**	1.1	-1.3	**-2.4**	1.1	**-2.1**	-1.3	**-2.0**

* KO_M, KO_T and KO_S refer to data of the *ccpA* knockout strain at the mid-exponential, the transition and the stationary stage of growth, respectively [154].

When focusing on the regulation of well-established CcpA-activated genes, the transcriptional effects are not straightforward as gene- and mutant-specific responses are observed (Table 3). For example, in the mid-exponential phase in the Y90P strain all eight genes are affected by the mutation and their expression ratios are in some cases even more pronounced than in the full knockout (see the ratios in grey in Table 3). Interestingly, in the stationary phase, the effects were not longer observed for all the genes. Contrary to Y90P, the T307W strain exhibited a very weak effect at both time points which means that this mutant has an activity comparable to the wild-type CcpA strain. In case of R54W and V301W, a mixed type of regulation takes place as two (*alsS* and *ilvB*) out of four genes showed differential expression in the mid-exponential phase. In the exponential phase both mutants seem to have a reduced activity. Remarkably, in both cases the regulation was present for *ackA* in the stationary phase as well. This phenomenon is in contradiction to the recently published time series data (ratio -1.1), however, under the experimental setup used in this study glucose was absent from the medium in the stationary phase [154]. In addition, previously proposed putative CcpA-activated genes [154], *ie*. *opuAA*, *opuE*, *yhbI* and *manA,* are also affected in one or more CcpA mutants analyzed in this paper.

Comparison of transcriptional and translation effects of the CcpA mutants.

The transcriptome data was compared to the translational data obtained by Sprehe *et al* [231]. In this study, the expression levels of four genes (see ß-gal in Table 4) were followed in several single amino acid mutants of CcpA in a *lacZ* fusion system. Although the ratios originating from the transcriptome analysis are in general less pronounced, the comparisons summarized in Table 4 reveal that, with some exceptions, the two approaches yield

qualitatively similar results despite the fact that Sprehe and colleagues used different media for the individual strains. The very high ratios in the case of the ß-galactosidase measurements (for example for *xynP* or *alsS*) result from the low ß-galactosidase activity of the wild-type CcpA strain.

Table 4. Comparison of transcriptome data with single-gene analysis data obtained with the *lacZ* reporter gene fusions. The reported values represent average fold induction (minus values) or fold repression (plus values) of mutant/wild-type obtained from either the microarrays (MA) or the ß-galactosidase activity measurements (ß-gal). For *xynP* and *gntR* the results were obtained in the mid-exponential phase whereas for *ackA* and *alsS* in the stationary phase. Since in the case of *gntR* the expression ratios from the microarrays were statistically not significant, the ratios of neighbouring gene (*gntK*) of the same operon were used instead.

gene	***xynP***		***gntR (gntK)***		***ackA***		***alsS***	
strain	MA	ß-gal	MA	ß-gal	MA	ß-gal	MA	ß-gal
R54W (+glu)	2.5	130.0	5.5	5.9	-3.9	-1.4	-12.6	-556.0
Y90P (+glu)	5.7	1.0	3.9	2.6	-1.4	-1.3	-15.2	1.0
V301W (+glu)	1.7	450.0	8.1	5.9	-2.7	1.1	-6.8	-15.4
T307W (+glu)	1.1	-1.4	-1.6	-1.9	-1.5	-1.1	-1.1	-1.0

Furthermore, some discrepancies are observed for the *xynP* gene (strains Y90P and T307W), but these may be to some extent explained by the fact that the measurements for both *xynP* and *gntR* were performed in a different medium (containing gluconate) than in the transcriptome experiments described here. In view of the other results presented in Table 4, it is also slightly surprising that no transcriptional effects were determined for *alsS* in the Y90P mutant by the *lacZ* fusion approach.

Genes with particular expression profiles in the CcpA mutants

From the overall analysis presented above it can be concluded that the CcpA mutants behave in two distinct manners: T307W resembles the wild-type CcpA, whereas the other three mutants were (almost) defective in CCR (CcpA knock-out-like phenotype). However, while looking more closely at the data, there are many genes with an exceptional expression pattern, which does not fit into the general CcpA transcriptome picture. Some examples of this are given in Tables 6 and 7. Although the impact of the T307W mutation was very limited on the well-established CcpA regulon members listed in DBTBS and the ones added in the paper of Lulko *et al* [154], there was a clear effect on other genes with a (predicted) *cre*-site (Table 5) or without it (data not shown). Only the data from the exponential phase is considered here, to exclude any possible effects due to different glucose concentrations in the media of the strains. Even though the expression ratios were lower for T307W when compared to the other three mutants, there is a clear indication that this mutant has a reduced capability of performing CCR on a limited number of genes some of which encode proteins involved in carbon metabolism, especially the TCA cycle (Table 5).

Table 5. Genes with particular expression in the T307W mutant during exponential growth. Location and *cre* sequences were taken from the DBTBS or PRODORIC databases.

Gene	R54W	Y90P	V301W	T307W	Loction*	*cre*	Function
amyE	-	4.2	-	2.2	-7	TGTAAGCGTTAACA	alpha-amylase
citB	1.7	5.3	5.4	2.2	+1260	AGGAAACGTTTAAA	aconitate hydratase
dctP	6.6	10.3	10.2	2.5	-43	TGAAAACGCTATCA	C4-dicarboxylate transport protein
glpF	7.2	4.6	6.7	2.6	-37	TTGACACCGCTTTCA	glycerol utilization
glpT	4.2	4.7	3.1	2.9	ND/IG	TGATAGCGCTTTCT	uptake of glycerol-3-phosphate
glpQ	3.0	3.9	2.2	2.5	ND/IG	TGATAGCGCTTTCT	glycerol metabolism
kdgR	2.1	1.7	2.5	2.6	-107	TTTGAAACCGATTTCA	regulation of the pectin utilization
odhA	5.6	4.8	8.1	2.8	ND/IG	TAAAAACGCTTCCA	2-oxoglutarate dehydrogenase subunit
odhB	5.2	5.3	8.3	4.1	ND/IG	TGAAAAAGTAACCA	2-oxoglutarate dehydrogenase subunit
sacA	12.4	26.3	23.1	2.7	ND	ND	sucrase-6-phosphate hydrolase
sacP	5.0	9.2	7.4	1.2	ND/IG	GATGAAAGCGTATTCTT	sucrose transport
sdhA	6.5	8.1	8.9	2.9	ND	ND	succinate dehydrogenase subunit
sdhB	4.5	7.4	7.5	2.8	ND	ND	succinate dehydrogenase subunit
sdhC	9.2	10.2	12.3	3.4	ND/ORF ND/IG	TGATACCGGTTACA TGAAAGCGCAGTCT	succinate dehydrogenase subunit
sucC	3.8	8.3	6.8	3.0	ND/ORF	AGTAAGCGTTTTAT	succinyl-CoA synthetase subunit
sucD	2.6	1.6	3.6	2.9	ND	ND	succinyl-CoA synthetase subunit

*Location of the start of *cre*-site with respect to the transcription start site (ND-not defined; IG-intergenic region, ORF-open-reading frame region). In **bold** genes listed in DBTBS

The *cre* sequences of these genes were dissimilar and present in the ORFs or in the intergenic regions, so neither the location of the *cre*-sites nor their composition revealed something particular. Also the other three mutants were not always consistently defective in CCR. As shown in Table 6, six target genes were still regulated by the CcpA mutated at position R54 or V301, two of them being known as CcpA-activated and four known as CcpA-repressed. Remarkably, four of these genes have two *cre*-sites present in either their upstream regions (*pta* and *ackA*) or in the ORFs (the *ara* operon and *hutP*). Besides, the most of *cre*-sites of these six genes share a common sequence TG(T/A)AAGCG, though this particular motif is also present in other genes (see for example *sucC*, Table 5).

Discussion

To the best of our knowledge, this work represents the first attempt to analyse the effects of CcpA single amino acid mutants on the genome wide transcription level. A limited number of publications is available which describe the effects of single amino acid exchanges in the CcpA sequence of *Bacillus megaterium* [138,139], *Lactobacillus casei* [62] and *B. subtilis* [231]. However, these papers described consequences of the introduced mutations on transcriptional activity of a maximum of four reporter genes measured by means of beta-galactosidase assays. In *B. megaterium* three specific point-mutants (at positions 4, 47, 49) showed a significantly reduced repression phenotype but their growth was not affected, an

effect that was opposite for a mutant at position 17 [138]. Interestingly, the results of corresponding mutations in *L. casei* were not similar [62], which indicates that the importance of a given residue in the CcpA sequence is species-specific.

Table 6. Genes with particular expression in the R54W, Y90P and V301W mutants during exponential growth.

gene	R54W	Y90P	V301W	T307W	location	*cre*	function
ackA	-	-3.5	-	-	-124 -74	TGTAAGCGTTCATC TGTAAGCGTTATCA	conversion of acetyl-CoA to acetate
pta	-	-2.1	-	-	-70 -11	TGAAAGCGCTATAA AGAAAGCGTTTTTG	phosphotransacetylase
acsA	-	2.1	1.9	-	+38	TGAAAGCGTTACCA	acetate utilization
amyE	-	4.2	-	2.2	-7	TGTAAGCGTTAACA	alpha-amylase
araAB	-	2.0	-	-	+50 +2171	TGAAAGCGTTTTAT TGAAAACGATTACA	arabinose transport and metabolism
hutP	1.8	3.6	-	-	+2 +203	GTTAATAGTTATCA TGAAACCGCTTCCA	regulator of *hutPHUIGM*
yvfH	2.3	-	3.9	-	ND/IG	TAGAAACCGCTTAC	similar to L-lactate permease
rocB	2.2	-	3.1	-	ND/IG	GTGGAGGCGCTGAC	arginine/ornithine utilization
drm	2.5	-	2.4	-	ND/IG	ATGAAAACGGTTTA	phosphopentomutase

*Location of the start *cre*-site with respect to the transcription start site (ND-not defined; IG-intergenic region). In **bold** genes listed in DBTBS. Predicted c*re*-sites of *yvfH*, *rocB* and *drm* are indicated in reference [154].

A single point mutation in the CcpA sequence may have different influences on activity of particular promoters as revealed by our approach. Just a mere example of this phenomenon is presented in Table 3 that shows regulation of the well-established CcpA-activated genes. Depending on the time point of analysis a particular mutation either may have hardly any effects (T307W) or may influence transcriptional regulation in an equivalent manner to the full knockout (Y90P). In addition a mixed effect may also take place, *ie*. a mutation affects only a part of the genes of the regulon (R54W and V301W). These observations strongly indicate that in many instances the mechanisms underlying CcpA regulation are very complex and can be gene-specific. It could be explained by the fact that several specific co-effectors, some of them of yet unknown nature, are present in the cytoplasm and that they play a role in a selective binding of CcpA to a given promoter region. It is also rather puzzling that the genes presented in Tables 5 and 6 show the transcriptional effects, which are opposite to the general trends emerging from the microarray analyses. Although no regular motif or other common features were found for these genes, many scenarios are feasible. As explained above most likely subtle modifications in affinity to different effectors or structural changes as well as interplay with other transcriptional regulators could account for the observed "divergent" transcriptome profiles.

The results of glucose utilisation showed that the wild-type strain had a very strong preference for glucose as the level of that carbohydrate began to decrease quickly after 3 hours of growth (Fig. 1). This behaviour was not observed for the R54W, Y90P and V301W

mutants since in case of these strains hardly any or very restricted reduction of the glucose concentration in the medium was observed after 7 hours of growth. Remarkably, the T307W strain demonstrated a preference for glucose uptake comparable to that of the wild-type situation. This behaviour of the T307W mutant correlates to the results obtained by the microarray approach. Apparently, this mutation has a less destructive effect on the CcpA structure and the protein has in principle a wild-type-like activity, at least on the genes responsible for preferential glucose consumption. In addition, the results of the other mutants, *ie*. reduced or no preference for glucose utilisation, further substantiate that they are phenotypically similar to the full knock-out, which again goes in line with the observations based on the transcriptome profiling. These point mutations interfere in the structure of the protein in such a way that it can no longer bind the target DNA sequences (R54W) or its interaction with the low molecular mass coeffectors (NADP, G6P or FBP) is possibly hampered. Again here, these observations are confirmed by the microarray results as R54W, Y90P and V301W exhibited transcriptomic profiles closely similar to the full knockout situation.

It has to be noted that the T307W mutant is not completely identical to the wild-type strain as a considerable number of genes were differentially regulated when this strain is compared to its wild-type counterpart. It can be speculated that the T307W mutation causes a slightly reduced affinity of CcpA to HPr-Ser-P which affects the expression of genes (for some examples see Table 5) that requires a strong HPr-Ser-P binding by CcpA (other possible explanations are put forward further in the text). On the other hand, the fact that this mutation is neutral and bears almost no consequences for CCR is rather surprising since the other two mutants (Y90P and V301W) showed an impaired CCR. As all three single amino acid exchanges are located in the HPr(Crh)-Ser-P binding region of CcpA, a less efficient or lack of co-repressor binding for these mutants is expected. Although the ability to bind co-repressor by the wild type CcpA has been shown to be required for the efficient complex formation with the DNA fragments *in vitro*, the CcpA mutants that exhibit Hpr-Ser-P-independent DNA binding and the permanent repression of xylose utilization genes were reported [134]. These kinds of mutations, including the T306I variant of *B. megaterium* that corresponds to the T307 residue in *B. subtilis*, apparently render the CcpA protein insensitive to co-repressor binding or they enforce a permanent DNA binding conformation of the regulator. Our data reveal that this permanent repression or activation of the T307W mutant is extended to many more genes than only *xylA*. It remains to be established whether this behavior can be attributed to an unaffected HPr(Crh)-Ser-P binding or relies on a conformational change of CcpA. It can be speculated that the T307W mutation reduces the interaction of CcpA with HPr to such a level that it only affects expression of a particular subset of genes (such as *opuAA* or *pta* in Table 3 and other genes listed in Table 5). On the other hand, independent of the influence of HPr-binding, the interaction with low-molecular weight effectors may be stimulated upon the mutation. It has been shown that CcpA-binding to certain promoters (*xylA*, *xynP*, *pta*, *glpF* for example) is very weak or not

detectable without co-effectors, whereas binding to other promoters is strong (*rocG*) [223]. This indicates that different genes may exhibit different sensitivities to CcpA variants with an enhanced co-factor binding and possibly reduced HPr-binding. These subtle changes in binding affinity of CcpA to its effectors *in vivo* in combination with the *cre*-site degeneracy results in such mixed transcriptome profiles of the mutants investigated here. Further experiments to investigate the affinity to the co-repressor as well as the DNA-binding abilities of the mutants tested here would be required to shed more light on the sophisticated and multiple mechanisms of the regulatory events performed by CcpA. It is also possible that in case of certain promoters other transcriptional factors take over regulation, hence masking the absence of functional CcpA. For example CodY, the same as CcpA, contributes to the negative regulation of *acsA*, *citB* and *hutP* in response to carbon availability and these genes showed unexpected behaviour in the mutants [69,124]. This could not be concluded for *ilvB* for which modes of regulation for CcpA and CodY are opposite [225].

It must not be forgotten that under our experimental set-up the Y90P and T307W strains showed at least two-times lower mRNA levels of *ccpA*. Under the assumption that this also results in decreased protein levels, the microarray outcome might be possibly influenced. However, taken into account a completely divergent behaviour of these two mutants, this is rather unlikely.

An interesting issue emerges from the array data presented here, since the genes that have been proved to belong to the CcpA regulon previously in other studies, but have shown not to be regulated (ratio 1) by the mutated variants of CcpA are also of interest. Lack of regulation in such cases points to the fact that an amino-acid exchange in a given position caused no damage in the structure of the regulator and as such did not alter its biological activity (wild-type situation). DNA microarray analysis is meant to look for differences in the mRNA levels which are statistically significant between two samples. Bearing that in mind, it is challenging to assess the reliability of expression ratios that are close to the value of one.

Moreover, it is possible that the wild-type CcpA and the mutants do not bind to DNA sequences of some genes under the tested conditions, thus differential expression cannot be observed in such cases. Nevertheless, the transcriptome data of the four CcpA mutants described here could serve as a starting point for future efforts to characterise different mechanisms through which CcpA exerts its array of functions *in vivo*. In this respect, it would be of great value to reassess the data in combination with the proteomics results of the same mutants (Glagla unpublished data) and concentrate further research on a selection of potentially interesting genes in order to shed more light on the complex mechanisms behind CcpA regulation.

Chapter 7

Summary and general discussion

Global view on the results

In this thesis a broad range of regulatory responses of *B. subtilis* to stress conditions have been investigated using DNA microarrays. Table 1 offers a bird's-eye view of the key results presented in the chapters of this thesis. Intricate regulation of transcription plays a pivotal role in the adaptation of bacteria to a constantly changing environment. A very suitable, if not the best, approach to investigate transcriptional changes upon the encountered circumstances is transcriptomics that allow measuring the expression of thousands of genes simultaneously. In 2001/2002 when this project began, the DNA microarray technology was relatively in its infancy and only a very limited number of publications was available at that time. The know-how on this technology was fairly restricted and therefore the setting-up of a DNA microarray facility back then was a challenging task (see Chapter 2).

Table 1. Short summary of the results presented in the experimental chapters of this thesis.

Chapter 2	Pitfalls of PCR amplification of *B. subtilis* genome and fabrication of amplicon DNA microarrays, optimization of experimental steps of the microarray procedure.
Chapter 3	Discovery of new putative members of the CssRS regulon: *citM*, *ylxF*, *yloA, ykoJ* and several genes of the *flgB* operon. Protein overproduction and secretion in the stationary phase of growth influence the processes of sporulation (inhibition) and motility (activation).
Chapter 4	Mild lactic acid stress (pH 5.5) induces several defence mechanisms that allow *B. subtilis* to grow unaffectedly mainly due to F0F1 ATPase activity, while severe lactic acid stress (pH 5.0) leads to an extensive transcriptional reprogramming, including activation of the SigB and SigW regulons.
Chapter 5	CcpA regulation expands throughout the whole life cycle. Identification of potential new member of CcpA regulon such as the *ycgN* and *ydh* operons (repression) and *opuE*, the *opuAABC*, *yhb* and *man* operons (activation).
Chapter 6	R54W, Y90P and V301W CcpA mutants have a severely reduced capability to perform carbon catabolite regulation whereas the T307 strain has a less impaired activity.

In particular the preparation of the whole set of amplicons, comprising more than 4000 genes of *B. subtilis*, proved to be a very challenging venture. But once the DNA microarray technology became fully operational, it offered a powerful tool to study various aspects of bacterial physiology and genetics, some examples of which are presented in chapters 3 to 6 of this thesis. In these chapters different array platforms were used:

(i) the commercial DNA macroarrays (on nylon membranes) in Chapter 3 on secretion stress response;

(ii) the amplicon microarrays based on the PCR products spotted in-house (Chapter 2 and Chapter 5);

(iii) and lastly, oligonucleotide arrays (~70-mers) spotted in-house, in Chapter 4 (lactic acid stress), Chapter 5 (time-resolved CcpA transcriptome analysis) and Chapter 6 (transcriptome analysis of CcpA point mutants).

Although, the initial batches of the amplicon microarrays gave good results (data not shown) due to technical problems with the reproducibility of their production, they were not further

implemented. Specifically, the overall signal intensity as well as the spot morphology became considerably deteriorated between the first and the consecutive spotting runs, which resulted in loss of a great part of the microarray data (Chapter 2). As a result of these problems, but also due to consistency reasons, the oligonucleotide-based microarray platform was employed in the studies described in the chapters 4, 5 and 6, since these microarrays gave an overall much better hybridization outcome as compared to the amplicon ones.

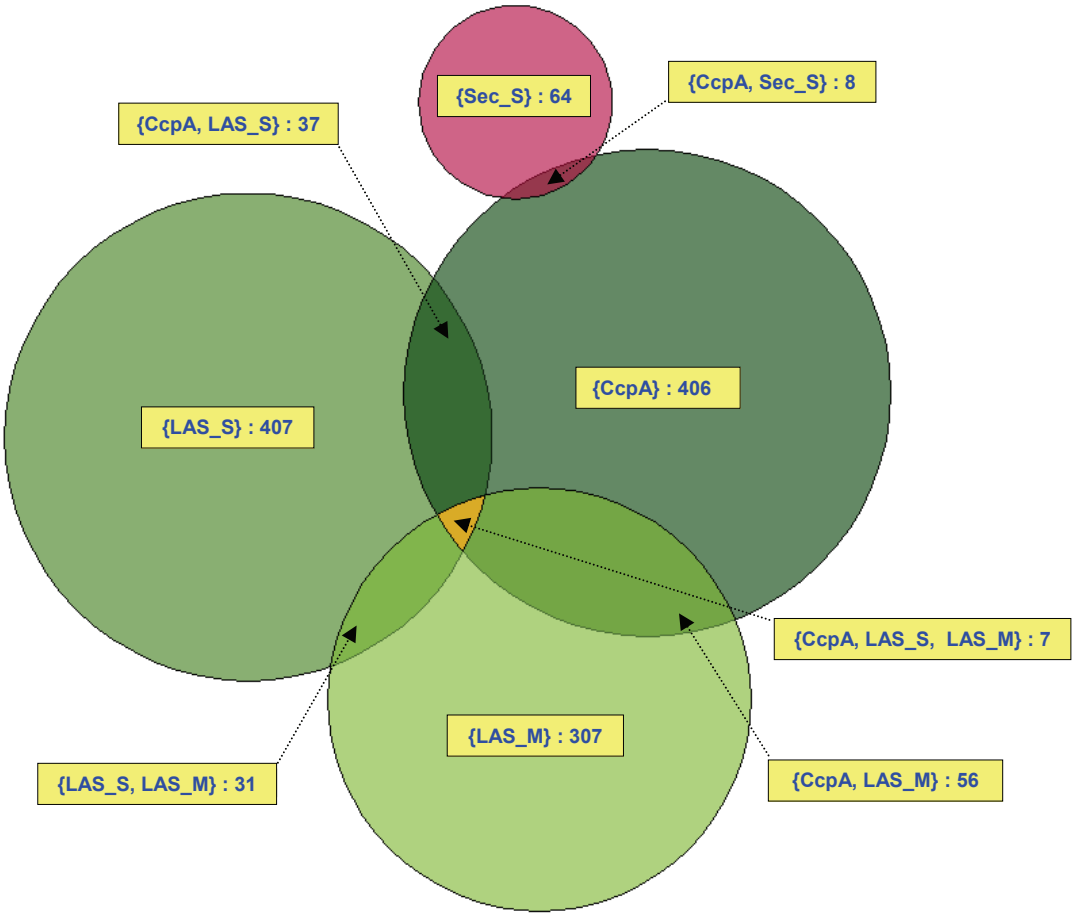


Figure 1. Overall relationships among the genes affected under the experimental conditions presented in this work expressed in the form of Venn diagrams. The sizes of circles and intersections reflect number of genes contained in each category (condition). The Venn mapping was performed on the genes that showed at least the 2-fold regulation. LAS_M and LAS_S correspond to mild and severe lactic acid stress. Sec_S – secretion stress.

The conditions explored in the experimental chapters of this thesis, i.e. the absence of a global transcriptional regulator, overproduction of a heterologous protein and exposure to a weak organic acid, are considered to be rather distinct. This is substantiated by the Venn diagram mapping performed with the VennMaster software (details of this kind of approach have been described previously by Kestler et al. [122]). The visualisation of the overall and

complex relationships among the expression profiles of the CcpA modulon (early till late exponential phase) and the secretion stress (late exponential and stationary phase) as well as the lactic acid stress stimulons (mild and severe) was achieved by means of Venn mapping. The areas of circles and the intersections are proportional to the numbers of differentially expressed genes of each condition or the numbers of genes common to two (or more conditions), respectively.

As revealed in Figure 1, despite a relatively high number of the affected genes, there is a limited number of shared genes among all the analyzed settings. While the overall outcome of this analysis indicates the divergent nature of the tested conditions, a modest overlap between the CcpA modulon and both lactic acid stress stimulons is present. This overlap results from a strong effect of the mild lactic acid stress on the CcpA regulon (approximately half of this regulon showed repression at 30 and 60 min after exposure to lactic acid at pH 5.5) as described in Chapter 4. A more detailed analysis of the 37 genes affected under severe lactic acid stress and the CcpA mutant experiments delivered no evidence for common functional categories (data not shown). This is not surprising since also attempts to identify significantly enriched functional categories in the dataset of the severe lactic stress alone was not very successful. Thus the observed patterns of gene regulation in response to the three investigated topics are distinct and should therefore be in principle treated as single entities.

Interpretation and relations of the results within the IOP project

The work presented in this thesis was a part of the comprehensive project titled "Comparative and predictive transcriptome analysis of Gram-positive bacteria for enhanced food functionality, quality and safety" (the IOP Genomics project coordinated by prof. O.P. Kuipers, grant IGE01018). One of the major goals within this project was the implementation of the DNA microarray technology to investigate the effects of mutations of global regulators (such as CcpA, see below) or various stresses on four evolutionary related Gram-positive bacteria:

(i) ***Bacillus subtilis* 168** (University of Groningen, results presented in this PhD thesis),

(ii) ***Lactococcus lactis*** MG1363 (University of Groningen, dr. A. Zomer, results presented in the PhD thesis, reference [289]),

(iii) ***Lactobacillus plantarum*** WCFS1 (University of Wageningen, dr. M. Stevens, results presented in the PhD thesis, reference [234]) and

(iv) ***Bacillus cereus*** ATCC14579 (University of Wageningen, dr. M. van der Voort, results presented in the PhD thesis, reference [249]).

Such an approach was aimed at not only unravelling relevant metabolic pathways and regulatory gene networks under the conditions met during industrial food fermentations but also at comparing and predicting the responses in these four organisms. The development of the high-throughput transcriptomic technologies was a prerequisite for the project and this task appeared to be more challenging and cost more time than originally expected (see

Chapter 2 for more details). Therefore, it was decided to restrict the focus of the project to two major topics, namely the transcriptional changes in the absence of the CcpA regulator (catabolite control protein A) and the response to exposure to lactic acid.

Based on the microarray analyses of the wildtype strains versus the *ccpA* mutants, which were performed at different stages of growth (from early exponential till stationary phase, see Fig. 1 in Chapter 5), several general conclusions for all the investigated bacteria could be drawn [154,234,250]:

(i) CcpA was shown to affect the expression of several hundreds of genes, which confirms its pleiotropic character.

(ii) CcpA has a predominant role as a repressor.

(iii) Functional categories of 'carbohydrate transport and metabolism' as well as 'general energy production and conversion' were commonly influenced in the four organisms.

(iv) In earlier phases of growth direct regulatory effects prevailed as *cre* sites were more frequently detected in the promoter regions of the affected genes, whereas in the later phases of growth the affected genes hardly contained *cre* sites, indicating the secondary nature of the observed effects.

(v) Time-resolved approach allowed for identification of new (putative) CcpA regulatory sites and possible new targets, at least in *B. subtilis* [154,234,250].

(vi) Thus in the stationary phase the observed effects were mostly indirect, which most likely resulted from the differences in media composition between the wild type and the mutant strains in the later phases of growth. For example, the glucose concentration was higher in the *B. subtilis* and *B. cereus* CcpA mutants and in case of the *L. plantarum* CcpA mutant acetate and lactate concentrations were different when compared to the wild type strain [154,234,250].

(vii) A link between CcpA-mediated regulation of carbon and nitrogen metabolism could not be established in *L. plantarum* and *B. cereus*. In *B. subtilis* control of the *gltAB* (glutamate synthase) and the *ilv-leu* (biosynthesis of branched-chain amino acids) operons by CcpA provides this link [153,258]. In *L. lactis* regulation of the peptidase gene *pepQ* by CcpA intertwines regulation of carbon and nitrogen metabolism [289].

(viii) Activation of glycolytic enzymes was not observed in *L. plantarum,* whereas it was in the other three bacteria.

Such a bird's eye view on the results reveals that several common general features concerning CcpA regulation can be indentified, but a deeper interpretation discloses diversity between the four microorganisms. Although a further and more detailed comparison of the obtained results is a rather complex and time-demanding task, it certainly gives an excellent opening for follow-up research.

Lactic acid as a stress factor has been chosen to gain better insights into the mechanisms governing resistance of microorganisms to weak organic acids, which is of great relevance

for the food industry, especially with regard to optimising production with probiotics or to determine the influence on survival of spoilage organisms during food preservation processes. The comparison of the transcriptome data obtained from the lactic acid stress experiments was also expected to deliver some similarities in response to the applied stress. On the contrary it occurred to be even less straightforward than the CcpA comparisons since the four bacteria responded differently with regard to their growth behaviour (see Table 2), which can be explained by the various natural habitats of these organisms. The initial idea was to fully synchronize the experimental set-ups, but due to the fundamental differences among these four bacteria it was not possible in practice, which was precisely expressed in their resistance level to lactic acid (see below). Yet, a common aspect of the experimental set-ups was that in all cases the samples for the analyses were collected at 2 to 4 time points at 3-60 minutes after the stress was applied. Several trends which emerged from the microarray analyses are highlighted in Table 2 and as can be seen the microarray methodology allowed to identify not only differences but also similarities among the strains. The importance of the experimental conditions is substantiated by the fact that mild (pH 5.5) and severe (pH 5.0) challenges triggered totally different transcriptional responses in *B. subtilis*. Likewise, in the case of *L. plantarum* the nature of the response strongly depended on the experimental approach. Lactic acid at a concentration of 100 mM at pH 5.0 (6.6 mM undissociated LA) in batch cultures was not harsh enough to induce a broad transcriptional reprogramming, which is reflected by the fact that under these conditions only 10 genes were affected [234]. Contrastingly, lactate anion at a concentration of 300 mM and slightly lower pH (pH 4.8, corresponding to ~31 mM undissociated LA) in the steady-state cultures evoked differential expression of multiple groups of genes [194]. Since the ratio between dissociated and undissociated forms of organic acids is pH-dependent, the broader range of effects observed at the lower pH values in *B. subtilis* and *L. plantarum* are attributed to the higher concentrations of undissociated, and regarded as more toxic, form of the acid.

A common observation is induction of the general and the oxidative stress response, indicating that upon lactic acid stress reactive oxygen moieties are formed, however the specific aspects of this response are generally quite different in the four bacteria discussed here. For example in *B. cereus* superoxide dismutase, catalases, and nitric oxide dioxygenase were induced. At pH 5.5 in *B. subtilis* a clear induction of several genes that belong to the CymR regulon was observed. Genes under control of CymR are involved in biosynthesis and recycling pathways of sulfur-containing amino acids and available literature data indicate the key role of this regulator in the control of cysteine levels [29,63,107]. The elevated expression of these genes likely lead to an increased pool of intracellular cysteine, the amino acid that is implicated in the oxidative stress response due to its antioxidant properties [147]. In addition, in *B. subtilis* YkuP flavodoxin, which supports nitric oxide (NO) synthesis from L-arginine [260], was transiently upregulated at pH 5.5. Thus, a possibly similar feature has been observed in both *Bacillus* species, namely, that the endogenous production of NO may provide a protection mechanism against acid stress.

Table 2. Response to lactic acid stress in four Gram-positive bacteria investigated within the IOP project.

	B. subtilis	*L. lactis*	*L. plantarum*	*B. cereus*
Experimental conditions	100mM LA (6.6 or 2.2 mM)* **pH 5.0 or pH 5.5** Batch culture	100mM LA (2.2 mM)* **pH 5.5** Batch culture	300mM LANa (30.7 mM)* **pH 4.8** Steady-state culture	0.7% (v/v) LA (2.0 mM)* **pH 5.5** Batch culture
Growth	Arrested / pH 5.0 Normal / pH 5.5	Severely impaired	Impaired	Arrested
General stress response	+	+	+	+
Oxidative stress	+	+	+	+
Carbohydrate metabolism	+	–	+	+
Amino acid metabolism	+	–	–	+
Nucleotide metabolism	–	+	0	–
Metabolic rerouting	0	0	+	+
ADI pathway	–	+		+ (?)
Lipid metabolism	+	+	+	–
Cell envelope modification	+	+ (?)	+	+/–
Multidrug transporters	+	0	0	–
F0F1 ATPase	+	0	0	–

*concentration of undissociated lactic acid at the indicated experimental pH
LA- lactic acid, LANa- sodium lactate
(**+**) upregulation, (**–**) downregulation, (**0**) no effect, (**?**) observed effect is of putative nature,
(**+/–**) up- or downregulation of genes was observed

Furthermore, genes of the Fnr regulon, including *ldh-lctP* (lactate dehydrogenase-permease) and cytochrome bd oxidase (*cydAB*), also seem to have a role in combating weak acid stress in *Bacillus* species. In *L. lactis* the genes encoding for putative DNA damage and SOS response as well as the super oxide dismutase had elevated mRNA levels and in *L. plantarum* the peroxide stress-related genes were induced [194]. In all four bacteria there was an effect on the cell-envelope (cell wall and/or membrane) and lipid metabolism, but the nature of this effect, similarly to the oxidative response, was again different in each organism.

In *B. subtilis* and *L. plantarum* elevated mRNA levels were observed for several members of the fatty acid biosynthetic pathway (Chapter 4, [194,241,289]). In *L. lactis* this effect was also seen, but it occurred transiently in a very early stage after exposure to lactic acid [194,289]. In *B. cereus* the murein hydrolase exporter and regulator genes were upregulated but other genes involved in cell envelope biogenesis were downregulated. Also expression of the genes of teichoic acid and capsular polysaccharide biosynthesis was modulated [180]. In *L. plantarum* multiple cell surface protein-encoding genes showed increased expression that resulted in morphological changes of the cell surface as proven by electron microscopy.

Furthermore, in this organism the genes involved in the biosynthesis of sterols were affected, which may lead to an increased rigidity of the membrane and a limited influx of lactic acid [194].

Discrepant expression patterns were noted for genes involved in carbohydrate, amino acid or nucleotide metabolism and other processes (see Table 2). Interestingly the genes encoding for the F0F1 ATPase showed a completely divergent response as well. In *B. subtilis* these genes were upregulated, in *L. lactis* and *L. plantarum* there was no effect on their expression, whereas in *B. cereus* their downregulation was observed. On the one hand this can be attributed to the different experimental conditions applied, but on the other hand it can also reflect the fact that in particular organisms the F0F1 ATPase activity is not necessarily regulated at gene expression level but at protein level. Metabolic rerouting of the pyruvate metabolism was indicated in two bacteria by induction of the butanediol fermentation pathway and part of the TCA cycle in *B. cereus*, whereas in *L. plantarum* by induction of the pentose phosphate cycle and pathways leading to generation of oxaloacetate and malate [180,194,289].

Last but not least, substantial numbers of genes with unknown function were differentially expressed in the four bacteria, which demonstrates that the cellular response and adaptation to lactic acid challenge is still not fully understood. However, it gives a window of opportunities for future research.

Concluding remarks and outlook for future

The research presented in this dissertation revolves around the DNA microarray technologies, which was the major tool exploited to obtain a transcriptomic "portray" of the paradigm Gram-positive bacterium, *B. subtilis*. Setting-up the basis of this methodology (Chapter 2) consumed an unexpected amount of effort and therefore most of the results presented in the subsequent four experimental chapters were generated in a rather short time-span. The performed transcriptome analyses have demonstrated their value as a screening method to accumulate large datasets that form a possible foundation for more detailed studies to further characterize various aspects of processes such as protein overproduction, catabolite regulation or weak acid response of *B. subtilis* and other related bacteria. Since DNA microarrays measure relative transcript abundance, but do not provide information on posttranscriptional or posttranslational regulation, the question stays open whether the corresponding changes also take place in the cell at protein and metabolite levels. Moreover, even when the expression level of a given gene and protein is correlated, it still does not indicate location nor activity and relationship with metabolites. In recent years it has become obvious that a single "omics" approach is not sufficient to characterize the complexity of biological systems [286]. For this reason expression profiling should be combined not only with other state of the art "omics" technologies, including proteomics and metabolomics, but also with classical biochemical and molecular biology methods. Such an integrated approach, which is a basis of systems biology [94], offers an opportunity to collect

more pieces of a jigsaw puzzle to build a comprehensive picture of cellular events, but even then it is illusive to expect that this jigsaw puzzle picture will be a complete one. Each individual technology has it own limitations, but these can be lessened by combination of different experimental approaches as well as striving to improve them. A prominent example of such a technological improvement is the development of the tiling DNA microarrays. These kinds of arrays not only cover both strands of the genome (thus not only one strand as in the case of "traditional" microarrays) but also include intergenic regions of a genome [80,205]. They are prepared with a set of overlapping oligonucleotide probes that span the entire genome at a high resolution [229,259]. The first results obtained with the tiling microarrays have been reported recently for *B. subtilis* and it was shown that through their implementation a discovery of new putative non-coding RNAs and antisense transcripts was possible [205]. Thus this kind of microarrays offers a higher potential than the traditional ones toward a comprehensive understanding of transcriptomes. Furthermore DNA microarrays suffer from several technical limitations, for example a relatively limited dynamic detection range of transcript levels due to saturation of the signal or high background levels owing to cross-hybridisations, and hence in the future they will give more and more space to other transcriptomics based methods. One of the possibilities is the high-throughput (next generation) sequencing of the cDNA converted from the whole cellular transcriptome content that allows investigating transcription at single nucleotide resolution without availability of a genome sequence [255]. However, since cDNA synthesis is not particularly suitable for the analysis of short, degraded and/or small quantity RNA samples as well as it may introduce multiple biases and artifacts, direct RNA sequencing without prior conversion of RNA to cDNA has potential to further revolutionize the field of transcriptomics [190,259]. Another important challenge to be tackled in future is determination of transcriptomes for single cells as the currently available tools measure the average level of gene expression or protein level in a mostly heterogeneous population consisting of millions or billions of cells. Single-cell transcriptomics is still in its infancy stage and is only operational in a eukaryotic setting [238], but once this approach becomes also available for prokaryotes it offers possibility to investigate the molecular basis for stochasticity in gene expression. This phenomenon is responsible for population heterogeneity that is one of the adaptation mechanisms upon confrontation with fluctuating environmental conditions [120,257].

Nederlandse samenvatting

Het werk beschreven in dit proefschrift richt zich op het onderzoek naar de regulatie van genexpressie van ***Bacillus subtilis*** in respons op stressomstandigheden met behulp van **DNA-microarrays**.

B. subtilis is een Gram-positieve, staafvormige bacterie die van nature voorkomt in de grond. Het organisme heeft verschillende eigenschappen waardoor het een interessant onderwerp is voor zowel de wetenschappelijke gemeenschap als de industrie. Deze eigenschappen omvatten onder andere de mogelijkheid tot:
(i) het opnemen van exogeen DNA (hetgeen genetische manipulaties vergemakkelijkt)
(ii) het vormen van endosporen (sporulatie; dit biedt een model voor het bestuderen van een relatief eenvoudige ontwikkelingsproces);
(iii) het uitscheiden van grote hoeveelheden eiwitten (waaronder heterologe eiwitten).

B. subtilis was de eerste Gram-positieve bacterie waarvoor de genoomsequentie is bepaald. Vanwege de grote hoeveelheid beschikbare informatie met betrekking tot de fysiologie en functionele genetica werd het een modelorganisme voor Gram-positieve bacteriën. Bovendien zijn er, dankzij de beschikbaarheid van de genoomsequentie, technologieën ontwikkeld die het monitoren van de expressie-niveaus van vrijwel alle genen tegelijk mogelijk maken (zogenaamde *transcriptomics*). Deze geavanceerde methoden boden een ongekende kans om te ontdekken hoe *B. subtilis* de meer dan 4100 genen van zijn genetische materiaal reguleert. Het is bekend dat er slechts een fractie van het totale aantal genen constitutief actief zijn (de zogenaamde *housekeeping* genen). De producten van deze genen zijn van cruciaal belang voor de cel onder alle omstandigheden (bijvoorbeeld enzymen van de glycolytische route, of ribosomale eiwitten). Aan de andere kant zijn veel andere eiwitten alleen 'op verzoek' nodig, in specifieke situaties zoals de aanwezigheid/afwezigheid van bepaalde stoffen, of onder stressomstandigheden. Teneinde de veranderingen in het milieu te kunnen herkennen en er adequaat op te kunnen reageren bezitten bacteriën moleculen, zogenaamde transcriptiefactoren, die de expressie van één of meer genen controleren. Zowel in natuurlijke omgevingen als tijdens industriële processen worden micro-organismen vaak geconfronteerd met ongunstige omstandigheden. Om hun overlevingskansen te vergroten hebben bacteriën mechanismen ontwikkeld ter confrontatie met en aanpassing aan de ondervonden stress. Over het algemeen bereiken ze dat door een herprogrammering van de genexpressie, gecoördineerd door de transcriptiefactoren, wat leidt tot synthese van eiwitten die gebruikt worden voor een snelle en optimale aanpassing aan een bepaalde stress. Een

dergelijke herprogrammering van de genexpressie kan worden bepaald en gekwantificeerd met behulp van DNA-microarrays.

DNA microarrays (ook wel DNA-chips genoemd) bevatten spots met DNA-fragmenten die individuele gensequenties vertegenwoordigen, veelal van één organisme. In theorie is het met microarrays mogelijk om de expressie-niveaus van alle mRNA-moleculen (*targets*) te bepalen die in een verzameling van cellen aanwezig zijn. Deze mRNA's coderen voor onder de gegeven groei-omstandigheden benodigde eiwitten, en zullen binden (hybridiseren) aan de complementaire DNA-strengen (*probes*) op de chip. Aangezien deze technologie het meten van veranderingen in de expressie-niveaus van duizenden genen in een enkel experiment mogelijk maakt, biedt zij een krachtig instrument voor het verkennen van de regulatie van genexpressie in bacteriën en andere organismen (zie figuur 4, hoofdstuk 2).

Er waren twee belangrijke doelstellingen van het project dat ten grondslag ligt aan dit proefschrift:
(i) de ontwikkeling van de transcriptomics technologie (**hoofdstuk 2**) en
(ii) de toepassing van deze technologie om de effecten van verschillende stresscondities, die zich bijvoorbeeld in het kader van industriële fermentatieprocessen voordoen, te onderzoeken.
De effecten van drie verschillende soorten stress zijn onderzocht:
(i) de overproductie van een heteroloog eiwit (**hoofdstuk 3**),
(ii) de blootstelling aan een zwak organisch zuur (**hoofdstuk 4**) en
(iii) de verstoring van het koolstof-metabolisme (**hoofdstuk 5 en 6**).

In **hoofdstuk 3** worden twee verschillende stammen van *B. subtilis* blootgesteld aan stress als resultaat van de overproductie van een heterologe α-amylase (AmyQ van *B. amyloliquefaciens*) op verschillende tijdspunten in de groei. Zoals hierboven uiteengezet, is *B. subtilis* in staat om grote hoeveelheden endogene eiwitten in het extracellulaire medium uit te scheiden. Daarom wordt deze bacterie vaak gebruikt als gastheer voor de productie en secretie van industrieel interessante heterologe enzymen. Echter, de secretie van heterologe eiwitten in grote hoeveelheden kan leiden tot verkeerd vouwen (*misfolding*) en afbraak van eiwitten. In *B. subtilis* wordt een ophoping van de verkeerd gevouwen eiwitten op de membraan/celwand-interface waargenomen door het CssRS twee-componenten-systeem, bestaande uit de in het membraan ingebedde sensor CssS en de transcriptie factor CssR. De producten van genen gereguleerd door CssRS fungeren ofwel als chaperones (die helpen bij het hervouwen van eiwitten of het ontvouwen ervan voorkomen) of als proteases (die helpen bij de afbraak van verkeerd gevouwen en geaggregeerde eiwitten). Zodoende fungeert het CssRS regulon als kwaliteitscontrole- en afweersysteem wanneer de cellen met secretie stress worden geconfronteerd.

In al onze array analyses werd de upregulatie van twee bekende doelgenen van het CssRS regulon, *htrA* en *htrB*, waargenomen, wat de kwaliteit van de verkregen transcriptoom-data valideert. Bovendien werden een aantal andere genen gevonden met een CssRS-afhankelijke veranderde transcriptie, bijvoorbeeld *citM, ylxF, yloA, ykoJ,* alsmede verschillende genen van het *flgB* operon. Een hoge-affiniteit CssR-binding *in vitro* werd echter alleen waargenomen voor *htrA* en *htrB*, en mogelijkerwijs *citM* (zie figuur 2, hoofdstuk 3). Daarnaast werd aangetoond dat er in de stationaire fase een aantal genen betrokken bij sporulatie lager, tot expressie komen, en een groep van motiliteit-specifieke (σD-afhankelijke) transcripten duidelijk hoger tot expressie kwamen door de overexpressie van AmyQ (zie figuur 1, hoofdstuk 3). Deze waarnemingen werden bevestigd met flowcytometrische analyses (zie figuur 3, hoofdstuk 3).

In **hoofdstuk 4** worden de groei en de globale genexpressie patronen in respons op blootstelling aan melkzuur bij verschillende zuurtegraden (pH's) geanalyseerd. Zwakke organische zuren zijn veelgebruikte conserveringsmiddelen, vanwege hun vermogen om bacteriële groei te remmen. Het verkrijgen van een beter inzicht in de mechanismen die de weerstand van micro-organismen tegen zwakke organische zuren regelen, is derhalve van groot belang voor de voedingsmiddelenindustrie. In deze studie hebben wij ons ten doel gesteld de afweermechanismen te ontrafelen die *Bacillus subtilis* in respons op melkzuur in werking stelt. Groei-experimenten in rijk, melkzuur bevattend, medium bij verschillende pH-niveaus hebben uitgewezen dat een subtiele pH-wijziging (van 5,2 naar 5,1) tot een plotselinge groeiremming leidt (zie figuur 1, hoofdstuk 4). Een dergelijk remming werd niet waargenomen bij cellen die blootgesteld waren aan zoutzuur, wat erop wijst dat dit groei-effect specifiek is voor melkzuur en niet kan worden toegeschreven aan verzuring van het medium alleen. De transcriptomics-experimenten werden uitgevoerd bij een pH van 5,5 en 5,0, aangezien de groei van *B. subtilis* bij deze pH-waarden respectievelijk niet beïnvloed (milde stress) of volledig gestaakt (ernstige stress) wordt. Uit de data bleek dat de eerste respons bij een pH van 5,5 (dat wil zeggen na 3 min) tot een duidelijke inductie leidde van genen onder controle van CymR, een transcriptie factor die de biosynthese en recyclingtrajecten van zwavelhoudende aminozuren reguleert. Bovendien werden verhoogde mRNA niveaus waargenomen van enkele genen, die voor de eiwitten met oxidoreductase functie coderen. Na 10 minuten werden andere afweermechanismen ingezet om de melkzuur-stress tegen te gaan; deze omvatten de inductie van enkele leden van het ijzer-opname regulator (Fur) regulon, de activiteit van de F0F1 ATPase en de up-regulatie van genen die voor (vermeende) multidrug transporters coderen (zoals *ebrB, yhcA* en *yubD*). Bovendien veroorzaakt blootstelling aan melkzuur (pH 5,5) waarschijnlijk een celmembraanstress op die in een differentiële expressie van vele, bij de vetstofwisseling betrokken genen resulteert. Bij zware melkzuur stress (pH 5,0) werd een uitgebreide transcriptionele herprogrammering waargenomen. Onder deze omstandigheden zijn het algemene (SigB-afhankelijke) stress regulon en het (SigW-afhankelijke) detoxificatie regulon

duidelijk geactiveerd (zie figuur 2, hoofdstuk 4). Op transcriptioneel niveau resulteerden een milde en een zware stress dus tot verschillende reacties. Teneinde de mogelijke rol van de geïdentificeerde genen in de weerstand tegen melkzuur te testen, werden vier kandidaten voor overexpressie geselecteerd. Van deze vier genen resulteerde alleen de overproductie van het membraaneiwit YoeB in een sneller groeiherstel na een zware melkzuur stress (zie figuur 4, hoofdstuk 4), mogelijk door een bijdrage van dit eiwit levert aan de stabilisatie van de cel-envelop. Over het algemeen leverde de overexpressie van de geselecteerde eiwitten echter geen sterk verbeterde weerstand tegen melkzuur.

In **hoofdstuk 5** werden het transcriptoom van *B. subtilis* 168 tijdens de groei in glucose-bevattend rijk medium vergeleken met een *ccpA* mutant. Het CcpA eiwit is de belangrijkste regulator van cataboliet controle in veel Gram-positieve bacteriën (zie figuur 1, hoofdstuk 1). Koolstof cataboliet repressie is een mechanisme waarmee cellen de beschikbare koolstofbronnen gebruiken in volgorde van voorkeur. In aanwezigheid van glucose wordt de expressie van genen voor het metabolisme van andere, moeilijker afbreekbare koolstofsubstraten, bijvoorbeeld uitgeschakeld. CcpA komt echter constitutief tot expressie, wat aangeeft dat deze regulator andere factoren nodig heeft om zijn pleiotrope functie te kunnen vervullen. CcpA vormt een dimeer en de DNA-bindende activiteit wordt gestimuleerd door de vorming van een complex met gefosforyleerd HPr (HPr-Ser-P) of het HPr-achtige eiwit Crh (Crh-Ser-P). Naast deze fosfoproteïnen kunnen ook laag moleculair gewicht moleculen, zoals NADP, glucose-6-fosfaat (G6P) en fructose-1,6-bisfosfaat (FBP), ofwel de DNA-bindende eigenschappen of de interactie van het CcpA-(HPr-Ser-P) complex met de transcriptie machinerie moduleren. CcpA reguleert de genexpressie door aan zogenaamde *cre*-elementen te binden, die zich zowel in de promoter regio's, of in coderende sequenties van de gereguleerde genen bevinden (zie figuur 1, hoofdstuk 1).
Hoewel de groei vergelijkbaar was, werd glucose onder onze experimentele omstandigheden volledig verbruikt door de wild-type stam in de stationaire fase, terwijl het nog steeds aanwezig was in het groeimedium van de mutante stam (zie figuur 1, hoofdstuk 5). Dientengevolge zijn er in dat stadium zowel directe als indirecte effecten op genexpressie waargenomen. Tijdens exponentiële groei beïnvloedt CcpA vooral de koolhydraten- en energiehuishouding, terwijl het eiwit vanaf de overgangsfase een breder scala aan fysiologische processen beïnvloedt, waaronder nucleotidemetabolisme, celbeweeglijkheid en eiwitsynthese (zie figuur 2, hoofdstuk 5). Een genoomwijde zoektocht bracht nieuwe *cre*-sites aan het licht, die volgens onze transcriptoom data *in vivo* kunnen functioneren. Een vergelijking van onze gegevens met de gepubliceerde gegevens van het transcriptoom van een *ccpA* mutant in de exponentiële groeifase bevestigde het eerder geïdentificeerde CcpA regulon (zie figuur 4, hoofdstuk 5). Daarnaast hebben we nieuwe genen geïdentificeerd, die mogelijk door CcpA worden onderdrukt (onder andere het *ycgN* en het *ydh* operon) of geactiveerd (zoals *opuE* en het *opuAABC* operon, de *yhb* en *man* operonen). Alle geactiveerde genen hebben een mogelijke *cre*-site, waarbij de positionering op het spiraalvormige DNA

van belang lijkt. Een vergelijkende analyse van deze genen met de bekende geactiveerde genen d.w.z. *ackA* en *pta,* bracht daarnaast de aanwezigheid van een mogelijke upstream activerende regio aan het licht, waarvan is aangetoond dat deze van belang is voor de activering van *ackA* (zie figuur 6, hoofdstuk 5). Over het geheel genomen suggereren deze microarray gegevens dat CcpA in een latere groeifase eveneens genexpressie kan reguleren, hetzij op zichzelf of gecomplexeerd met andere, nog onbekende, co-factoren.

In **hoofdstuk 6**, tot slot, werd transcriptoom analyse toegepast om de effecten te onderzoeken van puntmutaties in CcpA op gen-activering of repressie in de mid-exponentiële en stationaire groeifase van *Bacillus subtilis*. De gemuteerde aminozuurresiduen waren gelokaliseerd in het DNA-bindende domein (mutant residu R54W) of in het HPr (CrhP)-Ser-P bindende domein (mutant residuen Y90P, V301W en T307W). Uit de analyse is gebleken dat de T307W-mutatie een bescheiden invloed heeft op de werking van de CcpA, terwijl de overige drie mutaties, met name Y90P, de activiteit van deze transcriptionele regulator sterk beïnvloeden (zie figuur 2 en Tabel 2, hoofdstuk 6). De transcriptionele activiteiten van de mutanten werden op fenotypisch niveau bevestigd zoals blijkt uit de experimenten ten aanzien van het glucoseverbruik (zie figuur 1, hoofdstuk 6). De R54W, Y90P en V301W mutanten vertoonden geen voorkeur voor het gebruik van glucose, terwijl dit wel het geval was voor het wild type CcpA en de T307W mutant.

Streszczenie dla laików

Tematem tej pracy naukowej jest regulacja ekspresji genów u bakterii *Bacillus subtilis*. Do badania tych procesów posłużono się w głównej mierze nowoczesną technologią mikromacierzy DNA (ang. *DNA microarrays* lub *DNA chips*).

Bakterie

Bakterie to jedna z najprostszych form życia na ziemi. Są to mikroorganizmy występujące wszędzie wokół nas w niezliczonych formach i rodzajach. Zbudowane są z pojedynczej komórki, która sprawuje pełną kontrolę nad wszelkimi zachodzącymi w niej procesami. Bakterie posiadają niezwykłą umiejętność dostosowywania się do otoczenia, dzięki czemu mogą przeżyć nawet w najbardziej niekorzystnych dla nich środowiskach takich jak solanki, siarczane i termalne źródła. Nic dziwnego zatem, że są najczęściej spotykanym typem organizmu na naszej planecie. Występują zarówno głęboko w skorupie ziemskiej jak i wysoko w atmosferze. Można je spotkać na lodowcach Antarktydy, a nawet na terenach radioaktywnych. Doskonałym środowiskiem życia są dla nich również inne organizmy, czego prominentnym przykładem jest ciało człowieka. Naukowcy szacują, iż w ludzkich jelitach i na naszej skórze żyje ok. 500 do 1000 gatunków bakterii. Co ciekawe, w/na naszym ciele znajduje się co najmniej dziesięć razy więcej komórek bakteryjnych (ok. 100 trylionów) niż ludzkich (ok. 10 trylionów) o łącznej masie przekraczającej 1 kg!

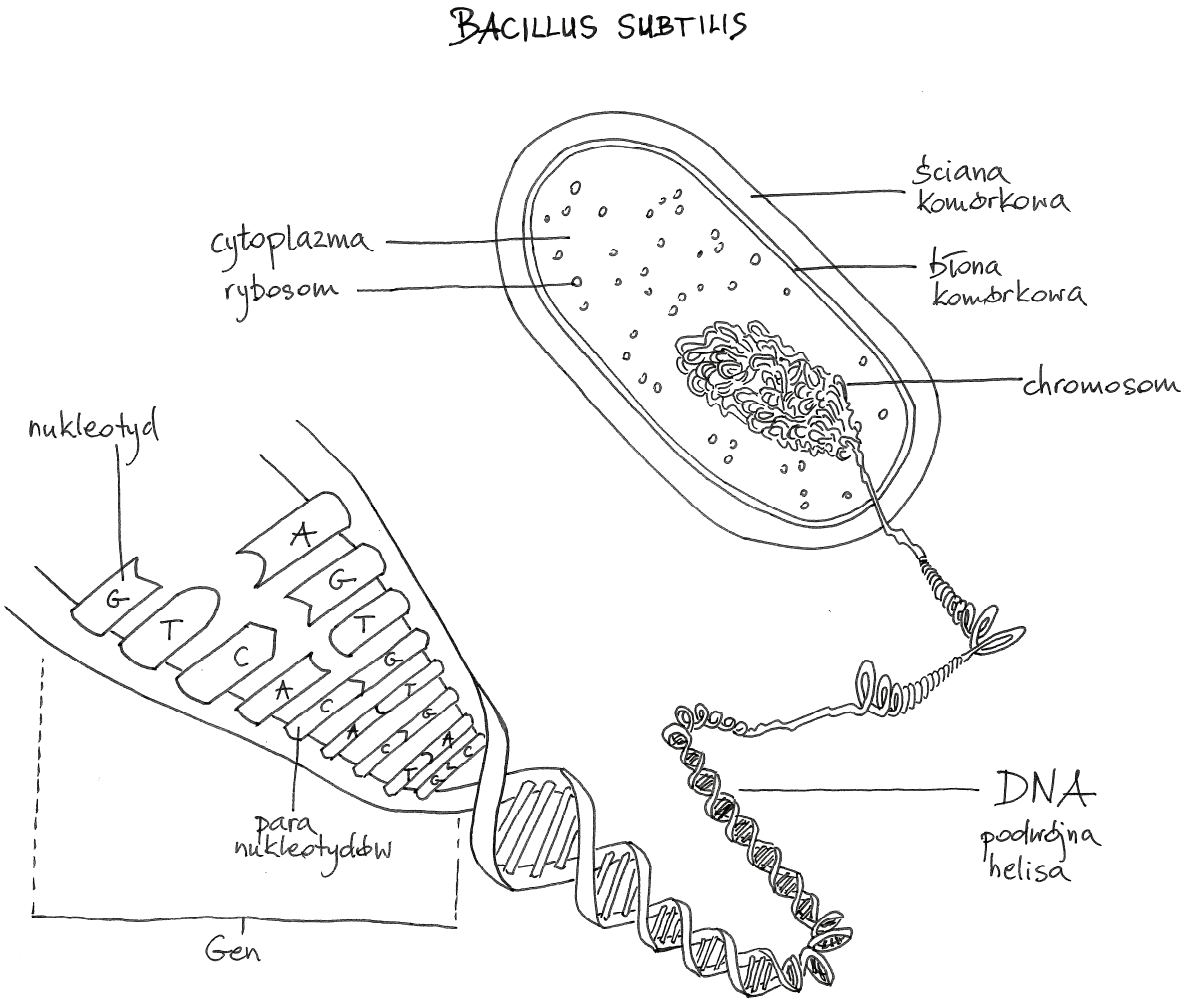
Bakterie można w uproszczeniu podzielić na pożyteczne i szkodliwe. Istnieją bakterie wywołujące choroby (np. *Salmonella*, *Listeria* czy bakteria wywołująca gruźlicę) czy powodujące psucie żywności. Nie wszystkie bakterie jednak są nieprzyjazne dla otoczenia. Większość z nich ma pozytywny wpływ na przebieg procesów biologicznych. Bakterie znajdujące się w jelitach wspomagają działanie układu trawiennego, a te na błonie śluzowej nosa czy oczu przyczyniają sie do utrzymania właściwego środowiska chroniąc tym samym nasz organizm przed wpływem szkodliwych drobnoustrojów. Inne bakterie natomiast wykorzystywane są do produkcji pożywienia np. jogurtów czy sera.

Właściwości bakterii, w tym możliwość dostosowywania się do zmian środowiska, podobnie jak cechy wszystkich innych organizmów zapisane są w ich DNA (skrót od: kwas deoksyrybonukleinowy) w postaci informacji genetycznej. Zbudowane z DNA geny są właściwym nośnikiem tej informacji. Niektóre organizmy posiadają zaledwie kilkadziesiąt genów, natomiast inne, bardziej złożone, mogą ich mieć ponad 30 000. Dla przykładu, ludzki

materiał genetyczny (tzw. genom) zawiera 20 000 – 25 000 genów kodujących białka. Dla porównania genom *B. subtilisa* ok. 4100 genów.

Bacillus subtilis

Bacillus subtilis jest to bakteria występująca powszechnie w glebie (Rys. 1). Potrafi ona z ogromną łatwością dostosowywać się do otoczenia i rozmnażać się w zmiennych warunkach środowiska jak zmiany temperatury, pH, zasolenia czy nawet brak dostępu do pokarmu.



Rys. 1. Schemat budowy DNA oraz komórki *B. subtilis*.

Ze względu na łatwość pozyskiwania i hodowli, a także szybkość z jaką się rozmnaża, *B. subtilis* chętnie jest wykorzystywany w przemyśle. Do jego popularności zarówno w przemyśle jak i wśród naukowców przyczynia się także wspólna dla wszystkich bakterii umiejętność wchłaniania DNA z otoczenia oraz możliwość produkcji białek (w tym też

białek innych organizmów w oparciu o ich informację genetyczną) i wydzielania ich na zewnątrz komórki. Pojedyncza komórka *B. subtilisa* może być zatem widziana jako swoista "fabryka białek". Biorąc pod uwagę fakt, że w 1/1000 litra pożywki bakteryjnej mogą znajdować się miliony komórek, bakteria ta jest niezwykle atrakcyjna nie tylko w przemyśle biotechnologicznym (hodowle bakteryjne sięgają rzędu nawet kilku tysięcy litrów) ale również farmaceutycznym i spożywczym.

Jak zbudowane są geny?

Nośnikiem informacji genetycznej są geny. Zbudowane są one z DNA - związku którego budulcem są nukleotydy: cztery podstawowe cegiełki oznaczone literami A, T, C i G (Rys. 1). Połączone w pary nukleotydy ułożone w określonej kolejności tworzą sekwencję genu. Kolejność ustawienia nukleotydów jest bardzo ważna, gdyż zawiera instrukcję wytwarzania białek. DNA ma postać podwójnej nici zwiniętej w spiralę zwaną helisą. Helisa DNA zwijając się dalej, przybiera kształt chromosomu. Kompletna informacja genetyczna nazywana jest genomem i składa się z określonej liczby chromosomów. Genom bakterii zbudowany jest z pojedynczego chromosomu. Dla porównania, genom ludzki składa się z 23 chromosomów.

Jak wygląda produkcja białek?

Wytwarzanie (zwane też ekspresją) białek odbywa się dzięki „przetłumaczniu" informacji genetycznej zawartej w DNA zgodnie z uproszczonym schematen. Geny ulegają ekspresji, gdy następuje „przepisanie" (tzw. transkrypcja) zawartej w nich informacji do RNA, na podstawie którego wytwarzane są białka (tzw. translacja czyli przetłumaczenie informacji z RNA na białko).

Produktami transkrypcji są jednoniciowe mRNA (informacyjne tzw. matrycowe RNA) będące pośrednikiem przenoszącym informację genetyczną między DNA a białkiem. Następnie w fazie translacji ma miejsce synteza białka. Odbywa się ona przy pomocy cząsteczek tRNA (transferowe RNA), które w oparciu o informację zawartą w mRNA dobierają odpowiednie aminokwasy, z których budowane jest białko. Im więcej kopii danego mRNA, tym więcej szans na wyższy poziom białka. Relatywną ilość wszystkich kopii RNA w komórce (tzw. profil ekspresji) można oszacować przy pomocy opisanej poniżej technologii mikromacierzy DNA.

Rola regulatorów

Zarządzanie tak dużą liczbą genów w tak małej przestrzeni, jaką jest komórka, to ogromne wyzwanie dla *B.subtilis*. Dlatego też ze względu na oszczędność energii ciągłej ekspresji podlegają jedynie nieliczne geny, których produkty są niezbędne do przeżycia bakterii w każdych okolicznościach (np. elementy rybosomów czyli maszynerii do syntezy białek w komórkach). Pozostałe geny natomiast angażowane są tylko "na żądanie" w określonych sytuacjach, np. pod wpływem stresu (wystawienia na niekorzystne warunki np. wysoka/niska temperatura czy zakwaszenie środowiska).
Organizmy posiadają czynniki transkrypcyjne, tzw. regulatory, czyli białka, które regulują aktywność genów na poziomie transkrypcji. Białka te wiążąc się do DNA, mogą hamować (represory) lub aktywować (aktywatory) proces ekspresji danego genu.

W sekwencji genomu *B. subtilisa* zakodowanych jest ponad 250 czynników transkrypcyjnych. Regulacja genów to bardzo złożone zjawisko ze względu na następujące aspekty:

(iv) regulatory mają możliwość występowania jako aktywatory lub represory jak również łączenia obu tych funkcji,
(v) jeden gen może być regulowany przez więcej niż jeden czynnik transkrypcyjny
(vi) jeden czynnik transkrypcyjny może modyfikować ekspresję innego.

Proces ten odgrywa niezwykle ważną rolę w dostosowywaniu się bakterii do otoczenia. Przykładem regulacji genów jest faworyzowanie glukozy jako najkorzystniejszego źródła energii w procesie zwanym **represją kataboliczną**. Bakterie zarządzają swoimi zasobami energii w jak najbardziej wydajny sposób zapewniający im jak najszybszy wzrost, a także pozwalający im przeżyć w konkurencyjnym środowisku. W procesie represji katabolicznej regulator o nazwie CcpA (ang. *catabolite control protein A*) hamuje ekspresję genów służących do pozyskiwania energii z bardziej złożonych cukrów, których rozkład wymaga większego nakładu energii w porównaniu z glukozą (CcpA występuje tu jako represor).

Mikromacierze DNA

Zasada działania

Wykorzystane w tej pracy mikromacierze DNA to czujnik (zwany też chipem) DNA umożliwiający ocenę profilu ekspresji w próbce zawierającej dany typ tkanki lub komórki bakteryjne. Mikromacierz DNA to niewielka płytka (np. plastikowa lub szklana), na którą naniesione są specjalnie przygotowane jednoniciowe cząsteczki DNA (tzw. sonda). Działanie mikromacierzy opiera się na podstawowej właściwości DNA polegającej na łączeniu się nukleotydów w pary zgodnie z zasadą, iż nukleotyd A tworzy parę tylko z T, a C łączy się tylko z G. Jest to tak zwana zasada komplementarności. Na przykład cząsteczka DNA o

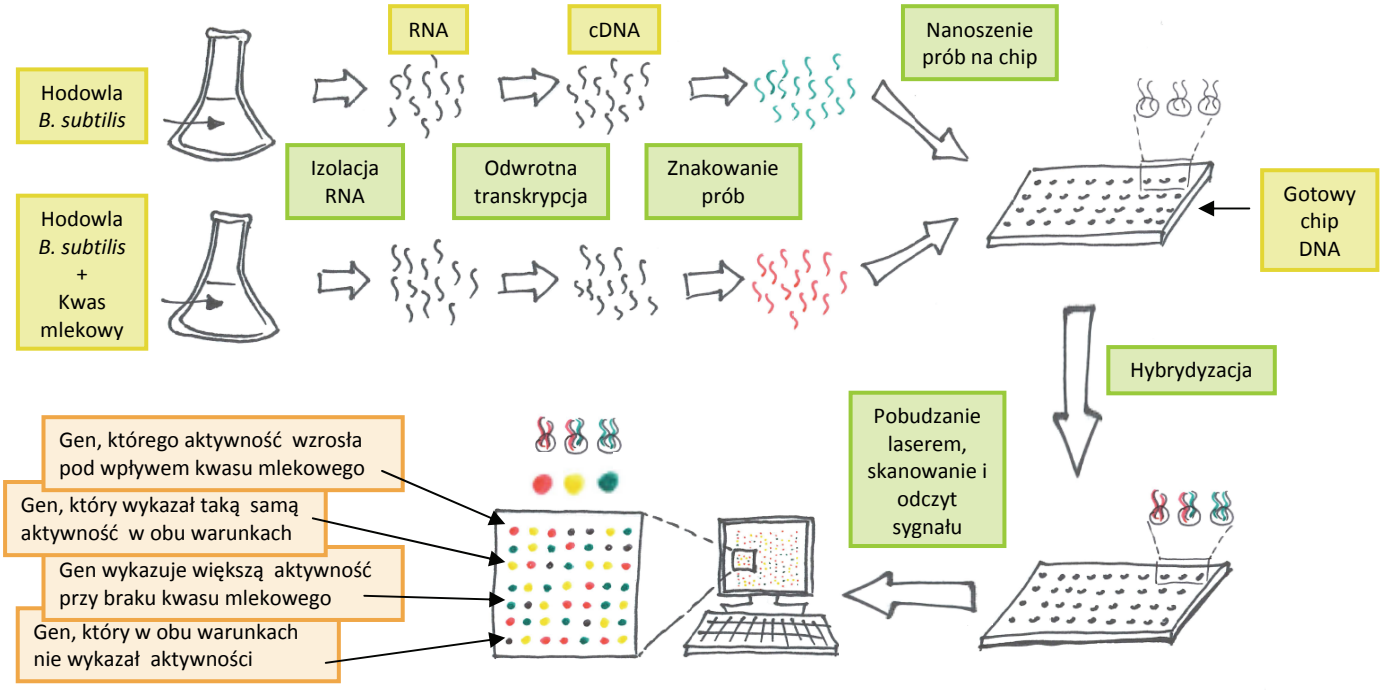
przykładowym fragmencie sekwencji TAGGCA połączy się ze znajdującą się na mikromacierzy sondą o sekwencji ATCCGT.

Wyniki testów i ich zastosowanie

Opisaną technikę stosuje się m.in. do wykrycia obecności ekspresji genów czy określenia ich sekwencji. Już przed wprowadzeniem mikromacierzy DNA na rynek (1996 rok) prowadzono badania oparte na zasadzie komplementarności DNA, które umożliwiały analizę kilku do kilkudziesięciu genów. Pojawienie się mikromacierzy DNA miało rewolucyjne znaczenie dla testów biologicznych, ponieważ umożliwiły one przetestowanie poziomu ekpresji wielu tysięcy genów, czy wręcz wszystkich genów danego organizmu, w jednym eksperymencie.

Przebieg eksperymentu

Do ekspermentów w niniejszym projekcie wykorzystano *B. subtilis* ze względu na pełną znajmość sekwencji genomu tej bakterii (scharakteryzowanego w 1997 roku). Każdy z ponad 4100 genów *B. subtilisa* znajduje się w ściśle określonym miejscu na przygotowanym wcześniej czujniku DNA. Wyizolowane z komórek *B. subtilis* cząsteczki mRNA poddane zostają procesowi odwrotnej transkrypcji, czyli przepisania informacji z mRNA z powrotem na DNA. Uzyskane w ten sposób twz. cDNA (komplementarne DNA) zostają oznaczone specjalnymi barwnikami fluororescencyjnymi (np. czerwonym i zielonym). Po naniesieniu prób zawierajacych wyznakowane cDNA na płytkę mikromacierzy następuje ich hybrydyzacja (t.j. wiązanie oparte na zasadzie komplementarnosci sekwencji) do sond DNA obecnych na płytce. Po hybrydyzacji mikromacierz pobudzona zostaje laserem w celu odczytu sygnału. Jesli dany gen jest aktywny (t.j. ulega ekspresji) w badanej próbce, odpowiadajaca mu sonda DNA bedzie emitować swiatło. Dane uzyskane z testów zostają komputerowo przekształcone w obraz, na którym intensywność poszczególnych kolorów pokazuje poziom ekspresji genu i jego zmiany wywołane zmianami warunków eksperymentalnych. Do oceny intensywności koloru i jej zamiany na liczby oraz dalszej analizy wyników niezbędne są różne specjalistyczne programy komputerowe. Eksperymenty wykorzystujące technologie mikromacierzy DNA generują olbrzymie ilości danych, dlatego też ich analiza jest dość trudnym i czasochłonnym wyzwaniem. Dla przykładu, w eksperymencie przedstawionym w rozdziale 5, w którym dwa różne szczepy *B. subtilisa* zostały przebadane w 4 niezależnych fazach wzrostu (próbki/komórki hodowano w cztery różne dni, tzw. powtórzenia biologiczne, a dodatkowo każda płytka DNA zawiera 2 kopie danego genu), zostało ostatecznie zebranych więcej niż ćwierć miliona pomiarów! Rys. 2 przedstawia przebieg eksperymentu na przykładzie ekspresji genów *B. subtilis* w środowisku neutralnym oraz w obecności kwasu mlekowego.



Rys. 2. Przebieg eksperymentu z użyciem mikromacierzy.

Zakres badań tej pracy

Niniejszy projekt zakładał realizację dwóch celów: (i) rozwój technologii mikromacierzy DNA oraz (ii) wdrożenie tej technologii w celu zbadania wpływu różnych warunków stresowych napotkanych na przykład w trakcie procesu fermentacji w przemyśle spożywczym. W pracy tej zanalizowano poziom ekspresji genów bakterii *B. subtilis* w odpowiedzi na stress czy zmienione warunki metabolizmu cukrów. Warunkiem wstępnym do realizacji projektu było rozwinięcie technologii mikromacierzy DNA, co okazało się bardzo trudnym i czasochłonnym wyzwaniem opisanym szczegółowo w **rozdziale 2**. Zaprezentowane badania w dalszych rozdziałach koncentrują się na trzech różnych czynnikach stresogennych:
(i) nadprodukcja białek heterogennych, czyli pochodzących z obcego organizmu (rozdział 3),
(ii) wystawienie na działanie słabego kwasu organicznego (rozdział 4) oraz
(iii) zaburzenia w regulacji metabolizmu węglowodanów (rozdział 5 i 6).

W **rozdziale 3** komórki *B. subtilis* poddane zostały stresowi wywołanemu przez nadprodukcję i wydzielanie heterogennej (tj. pochodzącej z innej bakterii) α-amylazy. Amylazy to enzymy wykorzystywane w przemyśle spożywczym. Ekspresja dziesiątek genów uległa zmianie, m.in. genów biorących udział w procesach sporulacji (tworzenia form

przetrwalnikowych) oraz poruszania się. Rozdział ten jest jedynym, na potrzeby którego zastosowano komercyjne makromacierze oraz metodę rodioaktywnego znakowania prób, z uwagi na fakt, iż metoda mikromacierzy nie była wówczas w pełni rozwinięta w naszym laboratorium.

Rozdział 4 prezentuje analizę wzrostu oraz profile ekspresji uzyskane po wystawieniu komórek na kwas mlekowy w dwóch różnych pH (miara zakwaszenia środowiska: im niższa wartość tym bardziej zakwaszone środowisko: pH 7.0 odpowiada neutralnemu środowisku; pH 5.5 w którym *B. subtilis* był w stanie się namnażać; oraz pH 5.0 w którym jego wzrost został calkowicie zahamowany). Słabe kwasy organiczne wykazują właściwości hamujące rozwój nieporządanych bakterii i są z tego względu często stosowanym środkiem konserwującym. Rozszyfrowanie mechanizmów regulujących odporność mikroorganizmów na te kwasy ma dlatego duże znaczenie dla przemysłu spożywczego. W przeprowadzonych tu eksperymentach zaobserowane zostały rozmaite mechanizmy adaptacji *B. subtilisa* w obecności kwasu mlekowego w pH 5.5, np. zmiany w ścianie i błonie komórkowej. Ponadto zidentyfikowane i zbadane zostały szczepy z nadekspresją genów w celu wskazania potencjalnych kandydatów charakteryzujących się ulepszonym wzrostem w warunkach silnego stresu kwasu mlekowego (pH 5.0). Tylko jeden z kandydatów wykazał lekką poprawę wzrostu w warunkach stresowych.

W **rozdziale 5** scharakteryzowany został regulon CcpA, czyli zestaw genów regulowanych przez ten czynnik transkrypcyjny. CcpA jest głównym regulatorem metabolizmu węglowodanów (cukrów). Eksperyment przeprowadzono w czterech różnych fazach wzrostu porównując komórki *B. subtilisa* posiadające regulator CcpA oraz te, z których został on usunięty (tzw. mutant CcpA). Takie podejście pozwoliło na pokazanie: (i) dynamicznego rozwoju regulonu CcpA w czasie, (ii) wpływu CcpA nie tylko na metabolizm cukrów, ale i też na różne inne procesy komórkowe oraz (iii) identyfikację nowych genów tego regulonu.

In **rozdziale 6** przeanalizowano wpływ mutacji punktowych (wymiana jednego aminokwasu) CcpA na jego funkcję jako regulator ekspresji genów. Analiza profilów ekspresji wykazała, iż jeden typ mutacji ma znikomy wpływ na działanie CcpA, podczas gdy pozostałe trzy w dużym stopniu wpływają na funkcjonowanie tego czynnika trankrypcyjnego. Wnioski te potwierdzone zostały przez eksperymenty ze zużyciem glukozy wykazujące, iż jedynie niezmutowany CcpA podobnie jak mutant z nieznacznie zmienionym profilem ekspresji w sposób preferencyjny zużywały zasoby glukozy w pożywce. Natomiast pozostałe trzy mutanty ze znacząco zmienionym profilem ekspresji nie zużywały tego cukru w preferencyjnej kolejności (podobnie jak i *B. subtilis* z całkowicie zmutownym CcpA), co podkreśla, iż te trzy mutacje punktowe mają „niszczący" wpływ na białko poprzez zmianę jego struktury.

Wreszcie **rozdział 7** prezentuje dyskusję na temat wszystkich danych uzyskanych w rozdziałach poprzedzających. Ponadto autor zajmuje krytyczne stanowisko wobec technologii mikromacierzy DNA oraz przedstawia perspektywy przyszłych badań.

References

1. Aerts S., Van L.P., Thijs G., Mayer H., de M.R., Moreau Y., and De M.B. (2005). TOUCAN 2: the all-inclusive open source workbench for regulatory sequence analysis. Nucleic Acids Res. *33*: W393-W396.
2. Aguilar P.S., Cronan J.E., Jr., and de M.D. (1998). A *Bacillus subtilis* gene induced by cold shock encodes a membrane phospholipid desaturase. J. Bacteriol. *180*: 2194-2200.
3. Albanesi D., Mansilla M.C., Schujman G.E., and de M.D. (2005). *Bacillus subtilis* cysteine synthetase is a global regulator of the expression of genes involved in sulfur assimilation. J. Bacteriol. *187*: 7631-7638.
4. Albano M., Smits W.K., Ho L.T., Kraigher B., Mandic-Mulec I., Kuipers O.P., and Dubnau D. (2005). The Rok protein of *Bacillus subtilis* represses genes for cell surface and extracellular functions. J. Bacteriol. *187*: 2010-2019.
5. Allenby N.E., O'Connor N., Pragai Z., Ward A.C., Wipat A., and Harwood C.R. (2005). Genome-wide transcriptional analysis of the phosphate starvation stimulon of *Bacillus subtilis*. J. Bacteriol. *187*: 8063-8080.
6. Anagnostopoulos C. and Spizizen J. (1961). Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. *81*: 741-746.
7. Antelmann H., Darmon E., Noone D., Veening J.W., Westers H., Bron S., Kuipers O.P., Devine K.M., Hecker M., and van Dijl J.M. (2003). The extracellular proteome of *Bacillus subtilis* under secretion stress conditions. Mol. Microbiol. *49*: 143-156.
8. Asai K., Yamaguchi H., Kang C.M., Yoshida K., Fujita Y., and Sadaie Y. (2003). DNA microarray analysis of *Bacillus subtilis* sigma factors of extracytoplasmic function family. FEMS Microbiol. Lett. *220*: 155-160.
9. Auger S., Danchin A., and Martin-Verstraete I. (2002). Global expression profile of *Bacillus subtilis* grown in the presence of sulfate or methionine. J. Bacteriol. *184*: 5179-5186.
10. Baerends R.J., Smits W.K., de Jong A., Hamoen L.W., Kok J., and Kuipers O.P. (2004). Genome2D: a visualization tool for the rapid analysis of bacterial transcriptome data. Genome Biol. *5*: R37.
11. Baichoo N., Wang T., Ye R., and Helmann J.D. (2002). Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. Mol. Microbiol. *45*: 1613-1629.
12. Baldi P. and Long A.D. (2001). A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. Bioinformatics. *17*: 509-519.
13. Barrett T. and Edgar R. (2006). Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. Methods Enzymol. *411*: 352-369.
14. Beckering C.L., Steil L., Weber M.H., Volker U., and Marahiel M.A. (2002). Genomewide transcriptional analysis of the cold shock response in *Bacillus subtilis*. J. Bacteriol. *184*: 6395-6402.
15. Belitsky B.R., Kim H.J., and Sonenshein A.L. (2004). CcpA-dependent regulation of *Bacillus subtilis* glutamate dehydrogenase gene expression. J. Bacteriol. *186*: 3392-3398.
16. Berka R.M., Hahn J., Albano M., Draskovic I., Persuh M., Cui X., Sloma A., Widner W., and Dubnau D. (2002). Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. Mol. Microbiol. *43*: 1331-1345.
17. Bisicchia P., Noone D., Lioliou E., Howell A., Quigley S., Jensen T., Jarmer H., and Devine K.M. (2007). The essential YycFG two-component system controls cell wall metabolism in *Bacillus subtilis*. Mol. Microbiol. *65*: 180-200.

18. Blencke H.M., Homuth G., Ludwig H., Mader U., Hecker M., and Stulke J. (2003). Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. Metab Eng *5*: 133-149.

19. Blencke H.M., Reif I., Commichau F.M., Detsch C., Wacker I., Ludwig H., and Stulke J. (2006). Regulation of *citB* expression in *Bacillus subtilis*: integration of multiple metabolic signals in the citrate pool and by the general nitrogen regulatory system. Arch. Microbiol. *185*: 136-146.

20. Blom E.J., Bosman D.W., van Hijum S.A., Breitling R., Tijsma L., Silvis R., Roerdink J.B., and Kuipers O.P. (2007). FIVA: Functional Information Viewer and Analyzer extracting biological knowledge from transcriptome data of prokaryotes. Bioinformatics.

21. Bolhuis A., Broekhuizen C.P., Sorokin A., van Roosmalen M.L., Venema G., Bron S., Quax W.J., and van Dijl J.M. (1998). SecDF of *Bacillus subtilis*, a molecular Siamese twin required for the efficient secretion of proteins. J. Biol. Chem. *273*: 21217-21224.

22. Bongers R.S., Veening J.W., Van W.M., Kuipers O.P., and Kleerebezem M. (2005). Development and characterization of a subtilin-regulated expression system in *Bacillus subtilis*: strict control of gene expression by addition of subtilin. Appl. Environ. Microbiol. *71*: 8818-8824.

23. Brazma A. (2009). Minimum Information About a Microarray Experiment (MIAME)-successes, failures, challenges. ScientificWorldJournal. *9*: 420-423.

24. Brazma A., Parkinson H., Sarkans U., Shojatalab M., Vilo J., Abeygunawardena N., Holloway E., Kapushesky M., Kemmeren P., Lara G.G. et al. (2003). ArrayExpress--a public repository for microarray gene expression data at the EBI. Nucleic Acids Res. *31*: 68-71.

25. Breitling R. and Herzyk P. (2005). Biological master games: using biologists' reasoning to guide algorithm development for integrated functional genomics. OMICS. *9*: 225-232.

26. Britton R.A., Eichenberger P., Gonzalez-Pastor J.E., Fawcett P., Monson R., Losick R., and Grossman A.D. (2002). Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of *Bacillus subtilis*. J. Bacteriol. *184*: 4881-4890.

27. Brul S. and Coote P. (1999). Preservative agents in foods. Mode of action and microbial resistance mechanisms. Int. J. Food Microbiol. *50*: 1-17.

28. Budde I., Steil L., Scharf C., Volker U., and Bremer E. (2006). Adaptation of *Bacillus subtilis* to growth at low temperature: a combined transcriptomic and proteomic appraisal. Microbiology *152*: 831-853.

29. Burguiere P., Fert J., Guillouard I., Auger S., Danchin A., and Martin-Verstraete I. (2005). Regulation of the *Bacillus subtilis ytmI* operon, involved in sulfur metabolism. J. Bacteriol. *187*: 6019-6030.

30. Caldwell R., Sapolsky R., Weyler W., Maile R.R., Causey S.C., and Ferrari E. (2001). Correlation between *Bacillus subtilis scoC* phenotype and gene expression determined using microarrays for transcriptome analysis. J. Bacteriol. *183*: 7329-7340.

31. Cao M. and Helmann J.D. (2004). The *Bacillus subtilis* extracytoplasmic-function sigmaX factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides. J. Bacteriol. *186*: 1136-1146.

32. Cao M., Kobel P.A., Morshedi M.M., Wu M.F., Paddon C., and Helmann J.D. (2002). Defining the *Bacillus subtilis* sigma(W) regulon: a comparative analysis of promoter consensus search, run-off transcription/macroarray analysis (ROMA), and transcriptional profiling approaches. J. Mol. Biol. *316*: 443-457.

33. Cao M., Salzberg L., Tsai C.S., Mascher T., Bonilla C., Wang T., Ye R.W., Marquez-Magana L., and Helmann J.D. (2003). Regulation of the *Bacillus subtilis* extracytoplasmic function protein sigma(Y) and its target promoters. J. Bacteriol. *185*: 4883-4890.

34. Cao M., Wang T., Ye R., and Helmann J.D. (2002). Antibiotics that inhibit cell wall biosynthesis induce expression of the *Bacillus subtilis* sigma(W) and sigma(M) regulons. Mol. Microbiol. *45*: 1267-1276.

35. Chillappagari S., Miethke M., Trip H., Kuipers O.P., and Marahiel M.A. (2009). Copper acquisition is mediated by YcnJ and regulated by YcnK and CsoR in *Bacillus subtilis*. J. Bacteriol. *191*: 2362-2370.

36. Choi S.K. and Saier M.H., Jr. (2005). Regulation of pho regulon gene expression by the carbon control protein A, CcpA, in *Bacillus subtilis*. J. Mol. Microbiol. Biotechnol. *10*: 40-50.

37. Choi S.K. and Saier M.H., Jr. (2005). Regulation of sigL expression by the catabolite control protein CcpA involves a roadblock mechanism in *Bacillus subtilis*: potential connection between carbon and nitrogen metabolism. J. Bacteriol. *187*: 6856-6861.

38. Choi S.Y., Reyes D., Leelakriangsak M., and Zuber P. (2006). The global regulator Spx functions in the control of organosulfur metabolism in *Bacillus subtilis*. J. Bacteriol. *188*: 5741-5751.

39. Chu F., Kearns D.B., Branda S.S., Kolter R., and Losick R. (2006). Targets of the master regulator of biofilm formation in *Bacillus subtilis*. Mol. Microbiol. *59*: 1216-1228.

40. Chuaqui R.F., Bonner R.F., Best C.J., Gillespie J.W., Flaig M.J., Hewitt S.M., Phillips J.L., Krizman D.B., Tangrea M.A., Ahram M. et al. (2002). Post-analysis follow-up and validation of microarray experiments. Nat. Genet. *32 Suppl*: 509-514.

41. Comella N. and Grossman A.D. (2005). Conservation of genes and processes controlled by the quorum response in bacteria: characterization of genes controlled by the quorum-sensing transcription factor ComA in *Bacillus subtilis*. Mol. Microbiol. *57*: 1159-1174.

42. Cotter P.D. and Hill C. (2003). Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiol. Mol. Biol. Rev. *67*: 429-53.

43. Crosa J.H. (1997). Signal transduction and transcriptional and posttranscriptional control of iron-regulated genes in bacteria. Microbiol. Mol. Biol. Rev. *61*: 319-336.

44. Cruz R.H., Hoffmann T., Marino M., Nedjari H., Presecan-Siedel E., Dreesen O., Glaser P., and Jahn D. (2000). Fermentative metabolism of *Bacillus subtilis*: physiology and regulation of gene expression. J. Bacteriol. *182*: 3072-3080.

45. Cunin R., Glansdorff N., Pierard A., and Stalon V. (1986). Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. *50*: 314-352.

46. Cybulski L.E., del S.G., Craig P.O., Espinosa M., and de M.D. (2004). *Bacillus subtilis* DesR functions as a phosphorylation-activated switch to control membrane lipid fluidity. J. Biol. Chem. *279*: 39340-39347.

47. Dallas P.B., Gottardo N.G., Firth M.J., Beesley A.H., Hoffmann K., Terry P.A., Freitas J.R., Boag J.M., Cummings A.J., and Kees U.R. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR - how well do they correlate? BMC. Genomics *6*: 59.

48. Darmon E., Noone D., Masson A., Bron S., Kuipers O.P., Devine K.M., and van Dijl J.M. (2002). A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. J. Bacteriol. *184*: 5661-5671.

49. Deutscher J., Francke C., and Postma P.W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol. Mol. Biol. Rev. *70*: 939-1031.

50. Deutscher J., Galinier A., and Martin-Verstraete I. (2002). Carbohydrate Uptake and Metabolism. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells.*, Sonenshein AL, Hoch JA, Losick R, eds. (Washington: American Society for Microbiology Press), pp. 129-150.

51. Deutscher J., Kuster E., Bergstedt U., Charrier V., and Hillen W. (1995). Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in gram-positive bacteria. Mol. Microbiol. *15*: 1049-1053.

52. Deutscher J., Reizer J., Fischer C., Galinier A., Saier M.H., Jr., and Steinmetz M. (1994). Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the *ptsH* gene confers catabolite repression resistance to several catabolic genes of *Bacillus subtilis*. J. Bacteriol. *176*: 3336-3344.

53. Deutscher J. and Saier M.H., Jr. (1983). ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. Proc. Natl. Acad. Sci. U. S. A *80*: 6790-6794.

54. Dharmadi Y. and Gonzalez R. (2004). DNA microarrays: experimental issues, data analysis, and application to bacterial systems. Biotechnol. Prog. *20*: 1309-1324.

55. Diehl F., Beckmann B., Kellner N., Hauser N.C., Diehl S., and Hoheisel J.D. (2002). Manufacturing DNA microarrays from unpurified PCR products. Nucleic Acids Res. *30*: e79.

56. Dubrac S. and Msadek T. (2008). Tearing down the wall: peptidoglycan metabolism and the WalK/WalR (YycG/YycF) essential two-component system. Adv. Exp. Med. Biol. *631*: 214-228.

57. Duy N.V., Mader U., Tran N.P., Cavin J.F., Tam l.T., Albrecht D., Hecker M., and Antelmann H. (2007). The proteome and transcriptome analysis of *Bacillus subtilis* in response to salicylic acid. Proteomics. *7*: 698-710.

58. Eiamphungporn W. and Helmann J.D. (2008). The *Bacillus subtilis* sigma(M) regulon and its contribution to cell envelope stress responses. Mol. Microbiol. *67*: 830-848.

59. Eichenberger P., Fujita M., Jensen S.T., Conlon E.M., Rudner D.Z., Wang S.T., Ferguson C., Haga K., Sato T., Liu J.S. et al. (2004). The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. PLoS. Biol. *2*: e328.

60. Ernst J. and Bar-Joseph Z. (2006). STEM: a tool for the analysis of short time series gene expression data. BMC. Bioinformatics. *7*: 191.

61. Escolar L., Perez-Martin J., and de L., V (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. J. Bacteriol. *181*: 6223-6229.

62. Esteban C.D., Mahr K., Monedero V., Hillen W., Perez-Martinez G., and Titgemeyer F. (2004). Complementation of a Delta *ccpA* mutant of *Lactobacillus casei* with CcpA mutants affected in the DNA- and cofactor-binding domains. Microbiology *150*: 613-620.

63. Even S., Burguiere P., Auger S., Soutourina O., Danchin A., and Martin-Verstraete I. (2006). Global control of cysteine metabolism by CymR in *Bacillus subtilis*. J. Bacteriol. *188*: 2184-2197.

64. Eymann C., Homuth G., Scharf C., and Hecker M. (2002). *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. J. Bacteriol. *184*: 2500-2520.

65. Faires N., Tobisch S., Bachem S., Martin-Verstraete I., Hecker M., and Stulke J. (1999). The catabolite control protein CcpA controls ammonium assimilation in *Bacillus subtilis*. J. Mol. Microbiol. Biotechnol. *1*: 141-148.

66. Farewell A., Kvint K., and Nystrom T. (1998). Negative regulation by RpoS: a case of sigma factor competition. Mol. Microbiol. *29*: 1039-1051.

67. Fawcett P., Eichenberger P., Losick R., and Youngman P. (2000). The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A *97*: 8063-8068.

68. Feucht A., Evans L., and Errington J. (2003). Identification of sporulation genes by genome-wide analysis of the sigmaE regulon of *Bacillus subtilis*. Microbiology *149*: 3023-3034.

69. Fisher S.H., Rohrer K., and Ferson A.E. (1996). Role of CodY in regulation of the *Bacillus subtilis* hut operon. J. Bacteriol. *178*: 3779-3784.

70. Fitzpatrick T.B., Amrhein N., and Macheroux P. (2003). Characterization of YqjM, an Old Yellow Enzyme homolog from *Bacillus subtilis* involved in the oxidative stress response. J. Biol. Chem. *278*: 19891-19897.

71. Fox S., Filichkin S., and Mockler T.C. (2009). Applications of ultra-high-throughput sequencing. Methods Mol. Biol. *553*: 79-108.

72. Fozo E.M., Kajfasz J.K., and Quivey R.G., Jr. (2004). Low pH-induced membrane fatty acid alterations in oral bacteria. FEMS Microbiol. Lett. *238*: 291-295.

73. Fozo E.M. and Quivey R.G., Jr. (2004). Shifts in the membrane fatty acid profile of *Streptococcus mutans* enhance survival in acidic environments. Appl. Environ. Microbiol. *70*: 929-936.

74. Fujita M. (2000). Temporal and selective association of multiple sigma factors with RNA polymerase during sporulation in *Bacillus subtilis*. Genes Cells *5*: 79-88.

75. Fujita Y., Miwa Y., Galinier A., and Deutscher J. (1995). Specific recognition of the *Bacillus subtilis gnt* cis-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr**.** Mol. Microbiol. *17***:** 953-960.

76. Gaballa A., Wang T., Ye R.W., and Helmann J.D. (2002). Functional analysis of the *Bacillus subtilis* Zur regulon**.** J. Bacteriol. *184***:** 6508-6514.

77. Galinier A., Haiech J., Kilhoffer M.C., Jaquinod M., Stulke J., Deutscher J., and Martin-Verstraete I. (1997). The *Bacillus subtilis crh* gene encodes a HPr-like protein involved in carbon catabolite repression**.** Proc. Natl. Acad. Sci. U. S. A *94***:** 8439-8444.

78. Gasson M.J. (1983). Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic *streptococci* after protoplast-induced curing**.** J. Bacteriol. *154***:** 1-9.

79. Gosseringer R., Kuster E., Galinier A., Deutscher J., and Hillen W. (1997). Cooperative and non-cooperative DNA binding modes of catabolite control protein CcpA from *Bacillus megaterium* result from sensing two different signals**.** J. Mol. Biol. *266***:** 665-676.

80. Gregory B.D. and Belostotsky D.A. (2009). Whole-genome microarrays: applications and technical issues**.** Methods Mol. Biol. *553***:** 39-56.

81. Grote A., Klein J., Retter I., Haddad I., Behling S., Bunk B., Biegler I., Yarmolinetz S., Jahn D., and Munch R. (2009). PRODORIC (release 2009): a database and tool platform for the analysis of gene regulation in prokaryotes**.** Nucleic Acids Res. *37***:** D61-D65.

82. Grundy F.J., Turinsky A.J., and Henkin T.M. (1994). Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilization genes by CcpA**.** J. Bacteriol. *176***:** 4527-4533.

83. Guedon E., Moore C.M., Que Q., Wang T., Ye R.W., and Helmann J.D. (2003). The global transcriptional response of *Bacillus subtilis* to manganese involves the MntR, Fur, TnrA and sigmaB regulons**.** Mol. Microbiol. *49***:** 1477-1491.

84. Gusnanto A., Calza S., and Pawitan Y. (2007). Identification of differentially expressed genes and false discovery rate in microarray studies**.** Curr. Opin. Lipidol. *18***:** 187-193.

85. Hall H.K. and Foster J.W. (1996). The role of fur in the acid tolerance response of *Salmonella typhimurium* is physiologically and genetically separable from its role in iron acquisition**.** J. Bacteriol. *178***:** 5683-5691.

86. Hamoen L.W., Smits W.K., de Jong A., Holsappel S., and Kuipers O.P. (2002). Improving the predictive value of the competence transcription factor (ComK) binding site in *Bacillus subtilis* using a genomic approach**.** Nucleic Acids Res. *30***:** 5517-5528.

87. Hamon M.A., Stanley N.R., Britton R.A., Grossman A.D., and Lazazzera B.A. (2004). Identification of AbrB-regulated genes involved in biofilm formation by *Bacillus subtilis***.** Mol. Microbiol. *52***:** 847-860.

88. Hanna M.N., Ferguson R.J., Li Y.H., and Cvitkovitch D.G. (2001). *uvrA* is an acid-inducible gene involved in the adaptive response to low pH in *Streptococcus mutans***.** J. Bacteriol. *183***:** 5964-5973.

89. Hansson M., Rutberg L., Schroder I., and Hederstedt L. (1991). The *Bacillus subtilis hemAXCDBL* gene cluster, which encodes enzymes of the biosynthetic pathway from glutamate to uroporphyrinogen III**.** J. Bacteriol. *173***:** 2590-2599.

90. Harwood C.R., Wipat A., and Pragai Z. (2003). Functional analysis of the *Bacillus subtilis* genome**.** Methods in Microbiology *33***:** 337-367.

91. Hayes C.S., Illades-Aguiar B., Casillas-Martinez L., and Setlow P. (1998). In vitro and in vivo oxidation of methionine residues in small, acid-soluble spore proteins from *Bacillus* species**.** J. Bacteriol. *180***:** 2694-2700.

92. Hecker M., Pane-Farre J., and Volker U. (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria**.** Annu. Rev. Microbiol. *61***:** 215-236.

93. Hecker M. and Volker U. (2001). General stress response of *Bacillus subtilis* and other bacteria**.** Adv. Microb. Physiol *44***:** 35-91.

94. Heinemann M. and Sauer U. (2010). Systems biology of microbial metabolism. Curr. Opin. Microbiol. *13*: 337-343.

95. Helmann J.D. (2002). The extracytoplasmic function (ECF) sigma factors. Adv. Microb. Physiol *46*: 47-110.

96. Helmann J.D., Marquez L.M., and Chamberlin M.J. (1988). Cloning, sequencing, and disruption of the *Bacillus subtilis sigma 28* gene. J. Bacteriol. *170*: 1568-1574.

97. Helmann J.D., Wu M.F., Kobel P.A., Gamo F.J., Wilson M., Morshedi M.M., Navre M., and Paddon C. (2001). Global transcriptional response of *Bacillus subtilis* to heat shock. J. Bacteriol. *183*: 7318-7328.

98. Henkin T.M., Grundy F.J., Nicholson W.L., and Chambliss G.H. (1991). Catabolite repression of alpha-amylase gene expression in *Bacillus subtilis* involves a trans-acting gene product homologous to the *Escherichia coli* lacl and galR repressors. Mol. Microbiol. *5*: 575-584.

99. Herbort M., Klein M., Manting E.H., Driessen A.J., and Freudl R. (1999). Temporal expression of the *Bacillus subtilis secA* gene, encoding a central component of the preprotein translocase. J. Bacteriol. *181*: 493-500.

100. Hicks K.A. and Grossman A.D. (1996). Altering the level and regulation of the major sigma subunit of RNA polymerase affects gene expression and development in *Bacillus subtilis*. Mol. Microbiol. *20*: 201-212.

101. Hoper D., Volker U., and Hecker M. (2005). Comprehensive characterization of the contribution of individual SigB-dependent general stress genes to stress resistance of *Bacillus subtilis*. J. Bacteriol. *187*: 2810-2826.

102. Huang X., Decatur A., Sorokin A., and Helmann J.D. (1997). The *Bacillus subtilis* sigma(X) protein is an extracytoplasmic function sigma factor contributing to survival at high temperature. J. Bacteriol. *179*: 2915-2921.

103. Huang X., Fredrick K.L., and Helmann J.D. (1998). Promoter recognition by *Bacillus subtilis* sigmaW: autoregulation and partial overlap with the sigmaX regulon. J. Bacteriol. *180*: 3765-3770.

104. Hueck C.J. and Hillen W. (1995). Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the gram-positive bacteria? Mol. Microbiol. *15*: 395-401.

105. Hueck C.J., Hillen W., and Saier M.H., Jr. (1994). Analysis of a cis-active sequence mediating catabolite repression in gram-positive bacteria. Res. Microbiol. *145*: 503-518.

106. Hueck C.J., Kraus A., Schmiedel D., and Hillen W. (1995). Cloning, expression and functional analyses of the catabolite control protein CcpA from *Bacillus megaterium*. Mol. Microbiol. *16*: 855-864.

107. Hullo M.F., Auger S., Soutourina O., Barzu O., Yvon M., Danchin A., and Martin-Verstraete I. (2007). Conversion of methionine to cysteine in *Bacillus subtilis* and its regulation. J. Bacteriol. *189*: 187-197.

108. Hyyrylainen H.L., Bolhuis A., Darmon E., Muukkonen L., Koski P., Vitikainen M., Sarvas M., Pragai Z., Bron S., van Dijl J.M. et al. (2001). A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. Mol. Microbiol. *41*: 1159-1172.

109. Hyyrylainen H.L., Sarvas M., and Kontinen V.P. (2005). Transcriptome analysis of the secretion stress response of *Bacillus subtilis*. Appl. Microbiol. Biotechnol. *67*: 389-396.

110. Inacio J.M., Costa C., and de Sa-Nogueira I. (2003). Distinct molecular mechanisms involved in carbon catabolite repression of the arabinose regulon in *Bacillus subtilis*. Microbiology *149*: 2345-2355.

111. Ireton K., Rudner D.Z., Siranosian K.J., and Grossman A.D. (1993). Integration of multiple developmental signals in *Bacillus subtilis* through the Spo0A transcription factor. Genes Dev. *7*: 283-294.

112. Jacob S., Allmansberger R., Gartner D., and Hillen W. (1991). Catabolite repression of the operon for xylose utilization from *Bacillus subtilis* W23 is mediated at the level of transcription and depends on a cis site in the *xylA* reading frame. Mol. Gen. Genet. *229*: 189-196.

113. Jarmer H., Berka R., Knudsen S., and Saxild H.H. (2002). Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. FEMS Microbiol. Lett. *206*: 197-200.

114. Jenal U. and Hengge-Aronis R. (2003). Regulation by proteolysis in bacterial cells. Curr. Opin. Microbiol. *6*: 163-172.

115. Jones B.E., Dossonnet V., Kuster E., Hillen W., Deutscher J., and Klevit R.E. (1997). Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. J. Biol. Chem. *272*: 26530-26535.

116. Jordan S., Hutchings M.I., and Mascher T. (2008). Cell envelope stress response in Gram-positive bacteria. FEMS Microbiol. Rev. *32*: 107-146.

117. Ju J., Mitchell T., Peters H., III, and Haldenwang W.G. (1999). Sigma factor displacement from RNA polymerase during *Bacillus subtilis* sporulation. J. Bacteriol. *181*: 4969-4977.

118. Jurgen B., Hanschke R., Sarvas M., Hecker M., and Schweder T. (2001). Proteome and transcriptome based analysis of *Bacillus subtilis* cells overproducing an insoluble heterologous protein. Appl. Microbiol. Biotechnol. *55*: 326-332.

119. Kaan T., Homuth G., Mader U., Bandow J., and Schweder T. (2002). Genome-wide transcriptional profiling of the *Bacillus subtilis* cold-shock response. Microbiology *148*: 3441-3455.

120. Kaern M., Elston T.C., Blake W.J., and Collins J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. *6*: 451-464.

121. Kempf B. and Bremer E. (1995). OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. J. Biol. Chem. *270*: 16701-16713.

122. Kestler H.A., Muller A., Gress T.M., and Buchholz M. (2005). Generalized Venn diagrams: a new method of visualizing complex genetic set relations. Bioinformatics. *21*: 1592-1595.

123. Khatri P. and Draghici S. (2005). Ontological analysis of gene expression data: current tools, limitations, and open problems. Bioinformatics. *21*: 3587-3595.

124. Kim H.J., Jourlin-Castelli C., Kim S.I., and Sonenshein A.L. (2002). Regulation of the *Bacillus subtilis ccpC* gene by CcpA and CcpC. Mol. Microbiol. *43*: 399-410.

125. Kim J.H., Guvener Z.T., Cho J.Y., Chung K.C., and Chambliss G.H. (1995). Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. J. Bacteriol. *177*: 5129-5134.

126. Kim J.H., Voskuil M.I., and Chambliss G.H. (1998). NADP, corepressor for the *Bacillus* catabolite control protein CcpA. Proc. Natl. Acad. Sci. U. S. A *95*: 9590-9595.

127. Kim J.H., Yang Y.K., and Chambliss G.H. (2005). Evidence that *Bacillus* catabolite control protein CcpA interacts with RNA polymerase to inhibit transcription. Mol. Microbiol. *56*: 155-162.

128. Kobayashi G., Toida J., Akamatsu T., Yamamoto H., Shida T., and Sekiguchi J. (2000). Accumulation of an artificial cell wall-binding lipase by *Bacillus subtilis wprA* and/or *sigD* mutants. FEMS Microbiol. Lett. *188*: 165-169.

129. Kobayashi K., Ogura M., Yamaguchi H., Yoshida K., Ogasawara N., Tanaka T., and Fujita Y. (2001). Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. J. Bacteriol. *183*: 7365-7370.

130. Koburger T., Weibezahn J., Bernhardt J., Homuth G., and Hecker M. (2005). Genome-wide mRNA profiling in glucose starved *Bacillus subtilis* cells. Mol. Genet. Genomics *274*: 1-12.

131. Kocabas P., Calik P., Calik G., and Ozdamar T.H. (2009). Microarray Studies in *Bacillus subtilis*. Biotechnol. J. *4*: 1012-1027.

132. Kohli R.M. and Massey V. (1998). The oxidative half-reaction of Old Yellow Enzyme. The role of tyrosine 196. J. Biol. Chem. *273*: 32763-32770.

133. Konings W.N., Albers S.V., Koning S., and Driessen A.J. (2002). The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments. Antonie Van Leeuwenhoek *81*: 61-72.

134. Kraus A., Kuster E., Wagner A., Hoffmann K., and Hillen W. (1998). Identification of a co-repressor binding site in catabolite control protein CcpA. Mol. Microbiol. *30*: 955-963.

135. Kruger S. and Hecker M. (1995). Regulation of the putative *bglPH* operon for aryl-beta-glucoside utilization in *Bacillus subtilis*. J. Bacteriol. *177*: 5590-5597.

136. Kuipers O.P., de Jong A., Baerends R.J., van Hijum S.A., Zomer A.L., Karsens H.A., den Hengst C.D., Kramer N.E., Buist G., and Kok J. (2002). Transcriptome analysis and related databases of *Lactococcus lactis*. Antonie Van Leeuwenhoek *82*: 113-122.

137. Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S. et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature *390*: 249-256.

138. Kuster E., Hilbich T., Dahl M.K., and Hillen W. (1999). Mutations in catabolite control protein CcpA separating growth effects from catabolite repression. J. Bacteriol. *181*: 4125-4128.

139. Kuster-Schock E., Wagner A., Volker U., and Hillen W. (1999). Mutations in catabolite control protein CcpA showing glucose-independent regulation in *Bacillus megaterium*. J. Bacteriol. *181*: 7634-7638.

140. Lashkari D.A., DeRisi J.L., McCusker J.H., Namath A.F., Gentile C., Hwang S.Y., Brown P.O., and Davis R.W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. Natl. Acad. Sci. U. S. A *94*: 13057-13062.

141. Leenhouts K.J., Kok J., and Venema G. (1989). Campbell-like integration of heterologous plasmid DNA into the chromosome of *Lactococcus lactis* subsp. lactis. Appl. Environ. Microbiol. *55*: 394-400.

142. Leichert L.I., Scharf C., and Hecker M. (2003). Global characterization of disulfide stress in *Bacillus subtilis*. J. Bacteriol. *185*: 1967-1975.

143. Leung Y.F. and Cavalieri D. (2003). Fundamentals of cDNA microarray data analysis. Trends Genet. *19*: 649-659.

144. Lewis P.J. and Marston A.L. (1999). GFP vectors for controlled expression and dual labelling of protein fusions in *Bacillus subtilis*. Gene *227*: 101-110.

145. Li Y.H., Hanna M.N., Svensater G., Ellen R.P., and Cvitkovitch D.G. (2001). Cell density modulates acid adaptation in *Streptococcus mutans*: implications for survival in biofilms. J. Bacteriol. *183*: 6875-6884.

146. Lin J.T., Connelly M.B., Amolo C., Otani S., and Yaver D.S. (2005). Global transcriptional response of *Bacillus subtilis* to treatment with subinhibitory concentrations of antibiotics that inhibit protein synthesis. Antimicrob. Agents Chemother. *49*: 1915-1926.

147. Lithgow J.K., Hayhurst E.J., Cohen G., Aharonowitz Y., and Foster S.J. (2004). Role of a cysteine synthase in *Staphylococcus aureus*. J. Bacteriol. *186*: 1579-1590.

148. Livak K.J. and Schmittgen T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*: 402-408.

149. Long A.D., Mangalam H.J., Chan B.Y., Tolleri L., Hatfield G.W., and Baldi P. (2001). Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. J. Biol. Chem. *276*: 19937-19944.

150. Lorca G.L., Chung Y.J., Barabote R.D., Weyler W., Schilling C.H., and Saier M.H., Jr. (2005). Catabolite repression and activation in *Bacillus subtilis*: dependency on CcpA, HPr, and HprK. J. Bacteriol. *187*: 7826-7839.

151. Lord M., Barilla D., and Yudkin M.D. (1999). Replacement of vegetative sigmaA by sporulation-specific sigmaF as a component of the RNA polymerase holoenzyme in sporulating *Bacillus subtilis*. J. Bacteriol. *181*: 2346-2350.

152. Ludwig H., Meinken C., Matin A., and Stulke J. (2002). Insufficient expression of the *ilv-leu* operon encoding enzymes of branched-chain amino acid biosynthesis limits growth of a *Bacillus subtilis ccpA* mutant. J. Bacteriol. *184*: 5174-5178.

153. Ludwig H., Rebhan N., Blencke H.M., Merzbacher M., and Stulke J. (2002). Control of the glycolytic gapA operon by the catabolite control protein A in *Bacillus subtilis*: a novel mechanism of CcpA-mediated regulation**.** Mol. Microbiol. *45***:** 543-553.

154. Lulko A.T., Buist G., Kok J., and Kuipers O.P. (2007). Transcriptome Analysis of Temporal Regulation of Carbon Metabolism by CcpA in *Bacillus subtilis* Reveals Additional Target Genes**.** J. Mol. Microbiol. Biotechnol. *12***:** 82-95.

155. Lulko A.T., Veening J.W., Buist G., Smits W.K., Blom E.J., Beekman A.C., Bron S., and Kuipers O.P. (2007). Production and secretion stress caused by overexpression of heterologous alpha-amylase leads to inhibition of sporulation and a prolonged motile phase in *Bacillus subtilis***.** Appl. Environ. Microbiol. *73***:** 5354-5362.

156. Lund P.A. (2001). Microbial molecular chaperones**.** Adv. Microb. Physiol *44***:** 93-140.

157. Lusetti S.L. and Cox M.M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks**.** Annu. Rev. Biochem. *71***:** 71-100.

158. Madec E., Laszkiewicz A., Iwanicki A., Obuchowski M., and Seror S. (2002). Characterization of a membrane-linked Ser/Thr protein kinase in *Bacillus subtilis*, implicated in developmental processes**.** Mol. Microbiol. *46***:** 571-586.

159. Mader U., Antelmann H., Buder T., Dahl M.K., Hecker M., and Homuth G. (2002). *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics**.** Mol. Genet. Genomics *268***:** 455-467.

160. Mader U., Homuth G., Scharf C., Buttner K., Bode R., and Hecker M. (2002). Transcriptome and proteome analysis of *Bacillus subtilis* gene expression modulated by amino acid availability**.** J. Bacteriol. *184***:** 4288-4295.

161. Majtan T., Bukovska G., and Timko J. (2004). DNA microarrays-techniques and applications in microbial systems**.** Folia Microbiol. (Praha) *49***:** 635-664.

162. Makita Y., Nakao M., Ogasawara N., and Nakai K. (2004). DBTBS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics**.** Nucleic Acids Res. *32***:** D75-D77.

163. Malik S., Zalenskaya K., and Goldfarb A. (1987). Competition between sigma factors for core RNA polymerase**.** Nucleic Acids Res. *15***:** 8521-8530.

164. Marquez L.M., Helmann J.D., Ferrari E., Parker H.M., Ordal G.W., and Chamberlin M.J. (1990). Studies of sigma D-dependent functions in *Bacillus subtilis***.** J. Bacteriol. *172***:** 3435-3443.

165. Marshall O.J. (2004). PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR**.** Bioinformatics. *20***:** 2471-2472.

166. Mascher T., Hachmann A.B., and Helmann J.D. (2007). Regulatory overlap and functional redundancy among *Bacillus subtilis* extracytoplasmic function sigma factors**.** J. Bacteriol. *189***:** 6919-6927.

167. Mascher T., Margulis N.G., Wang T., Ye R.W., and Helmann J.D. (2003). Cell wall stress responses in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon**.** Mol. Microbiol. *50***:** 1591-1604.

168. Miethke M., Westers H., Blom E.J., Kuipers O.P., and Marahiel M.A. (2006). Iron starvation triggers the stringent response and induces amino acid biosynthesis for bacillibactin production in *Bacillus subtilis***.** J. Bacteriol. *188***:** 8655-8657.

169. Mirel D.B. and Chamberlin M.J. (1989). The *Bacillus subtilis* flagellin gene (*hag*) is transcribed by the sigma 28 form of RNA polymerase**.** J. Bacteriol. *171***:** 3095-3101.

170. Mirel D.B., Estacio W.F., Mathieu M., Olmsted E., Ramirez J., and Marquez-Magana L.M. (2000). Environmental regulation of *Bacillus subtilis* sigma(D)-dependent gene expression**.** J. Bacteriol. *182***:** 3055-3062.

171. Miwa Y. and Fujita Y. (2001). Involvement of two distinct catabolite-responsive elements in catabolite repression of the *Bacillus subtilis* myo-inositol (*iol*) operon**.** J. Bacteriol. *183***:** 5877-5884.

172. Miwa Y., Nagura K., Eguchi S., Fukuda H., Deutscher J., and Fujita Y. (1997). Catabolite repression of the *Bacillus subtilis gnt* operon exerted by two catabolite-responsive elements**.** Mol. Microbiol. *23***:** 1203-1213.

173. Miwa Y., Nakata A., Ogiwara A., Yamamoto M., and Fujita Y. (2000). Evaluation and characterization of catabolite-responsive elements (*cre*) of *Bacillus subtilis*. Nucleic Acids Res. *28*: 1206-1210.

174. Miwa Y., Saikawa M., and Fujita Y. (1994). Possible function and some properties of the CcpA protein of *Bacillus subtilis*. Microbiology *140 (Pt 10)*: 2567-2575.

175. Moch C., Schrogel O., and Allmansberger R. (2000). Transcription of the *nfrA-ywcH* operon from *Bacillus subtilis* is specifically induced in response to heat. J. Bacteriol. *182*: 4384-4393.

176. Moir-Blais T.R., Grundy F.J., and Henkin T.M. (2001). Transcriptional activation of the *Bacillus subtilis ackA* promoter requires sequences upstream of the CcpA binding site. J. Bacteriol. *183*: 2389-2393.

177. Moliere N. and Turgay K. (2009). Chaperone-protease systems in regulation and protein quality control in *Bacillus subtilis*. Res. Microbiol. *160*: 637-644.

178. Molle V., Fujita M., Jensen S.T., Eichenberger P., Gonzalez-Pastor J.E., Liu J.S., and Losick R. (2003). The Spo0A regulon of *Bacillus subtilis*. Mol. Microbiol. *50*: 1683-1701.

179. Molle V., Nakaura Y., Shivers R.P., Yamaguchi H., Losick R., Fujita Y., and Sonenshein A.L. (2003). Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. J. Bacteriol. *185*: 1911-1922.

180. Mols M., van K.R., Tempelaars M.H., van S.W., Moezelaar R., and Abee T. (2010). Comparative analysis of transcriptional and physiological responses of *Bacillus cereus* to organic and inorganic acid shocks. Int. J. Food Microbiol. *137*: 13-21.

181. Moreno M.S., Schneider B.L., Maile R.R., Weyler W., and Saier M.H., Jr. (2001). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. Mol. Microbiol. *39*: 1366-1381.

182. Mostertz J., Hochgrafe F., Jurgen B., Schweder T., and Hecker M. (2008). The role of thioredoxin TrxA in *Bacillus subtilis*: a proteomics and transcriptomics approach. Proteomics. *8*: 2676-2690.

183. Msadek T. (1999). When the going gets tough: survival strategies and environmental signaling networks in *Bacillus subtilis*. Trends Microbiol. *7*: 201-207.

184. Nguyen T.T., Eiamphungporn W., Mader U., Liebeke M., Lalk M., Hecker M., Helmann J.D., and Antelmann H. (2009). Genome-wide responses to carbonyl electrophiles in *Bacillus subtilis*: control of the thiol-dependent formaldehyde dehydrogenase AdhA and cysteine proteinase YraA by the MerR-family regulator YraB (AdhR). Mol. Microbiol. *71*: 876-894.

185. Nguyen V.D., Wolf C., Mader U., Lalk M., Langer P., Lindequist U., Hecker M., and Antelmann H. (2007). Transcriptome and proteome analyses in response to 2-methylhydroquinone and 6-brom-2-vinyl-chroman-4-on reveal different degradation systems involved in the catabolism of aromatic compounds in *Bacillus subtilis*. Proteomics. *7*: 1391-1408.

186. Nijland R., Heerlien R., Hamoen L.W., and Kuipers O.P. (2007). Changing a single amino acid in *Clostridium perfringens* {beta}-toxin affects efficiency of heterologous secretion by *Bacillus subtilis*. Appl. Environ. Microbiol.

187. Noone D., Howell A., and Devine K.M. (2000). Expression of *ykdA*, encoding a *Bacillus subtilis* homologue of HtrA, is heat shock inducible and negatively autoregulated. J. Bacteriol. *182*: 1592-1599.

188. Ogura M., Yamaguchi H., Kobayashi K., Ogasawara N., Fujita Y., and Tanaka T. (2002). Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. J. Bacteriol. *184*: 2344-2351.

189. Oomes S.J., Jonker M.J., Wittink F.R., Hehenkamp J.O., Breit T.M., and Brul S. (2009). The effect of calcium on the transcriptome of sporulating *B. subtilis* cells. Int. J. Food Microbiol. *133*: 234-242.

190. Ozsolak F., Platt A.R., Jones D.R., Reifenberger J.G., Sass L.E., McInerney P., Thompson J.F., Bowers J., Jarosz M., and Milos P.M. (2009). Direct RNA sequencing. Nature *461*: 814-818.

191. Padan E., Zilberstein D., and Schuldiner S. (1981). pH homeostasis in bacteria. Biochim. Biophys. Acta *650*: 151-166.

192. Palva I. (1982). Molecular cloning of alpha-amylase gene from *Bacillus amyloliquefaciens* and its expression in *B. subtilis*. Gene *19*: 81-87.

193. Petersohn A., Brigulla M., Haas S., Hoheisel J.D., Volker U., and Hecker M. (2001). Global analysis of the general stress response of *Bacillus subtilis*. J. Bacteriol. *183*: 5617-5631.

194. Pieterse B., Leer R.J., Schuren F.H., and van der Werf M.J. (2005). Unravelling the multiple effects of lactic acid stress on *Lactobacillus plantarum* by transcription profiling. Microbiology *151*: 3881-3894.

195. Piggot P.J. and Hilbert D.W. (2004). Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol. *7*: 579-586.

196. Poncet S., Mijakovic I., Nessler S., Gueguen-Chaignon V., Chaptal V., Galinier A., Boel G., Maze A., and Deutscher J. (2004). HPr kinase/phosphorylase, a Walker motif A-containing bifunctional sensor enzyme controlling catabolite repression in Gram-positive bacteria. Biochim. Biophys. Acta *1697*: 123-135.

197. Postma P.W., Lengeler J.W., and Jacobson G.R. (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. *57*: 543-594.

198. Presecan-Siedel E., Galinier A., Longin R., Deutscher J., Danchin A., Glaser P., and Martin-Verstraete I. (1999). Catabolite regulation of the *pta* gene as part of carbon flow pathways in *Bacillus subtilis*. J. Bacteriol. *181*: 6889-6897.

199. Price C.W., Fawcett P., Ceremonie H., Su N., Murphy C.K., and Youngman P. (2001). Genome-wide analysis of the general stress response in *Bacillus subtilis*. Mol. Microbiol. *41*: 757-774.

200. Priest F.G. (1977). Extracellular enzyme synthesis in the genus *Bacillus*. Bacteriol. Rev. *41*: 711-753.

201. Puri-Taneja A., Paul S., Chen Y., and Hulett F.M. (2006). CcpA causes repression of the *phoPR* promoter through a novel transcription start site, P(A6). J. Bacteriol. *188*: 1266-1278.

202. Qiu J. and Helmann J.D. (2001). The -10 region is a key promoter specificity determinant for the *Bacillus subtilis* extracytoplasmic-function sigma factors sigma(X) and sigma(W). J. Bacteriol. *183*: 1921-1927.

203. Quackenbush J. (2001). Computational analysis of microarray data. Nat. Rev. Genet. *2*: 418-427.

204. Quackenbush J. (2002). Microarray data normalization and transformation. Nat. Genet. *32 Suppl*: 496-501.

205. Rasmussen S., Nielsen H.B., and Jarmer H. (2009). The transcriptionally active regions in the genome of *Bacillus subtilis*. Mol. Microbiol. *73*: 1043-1057.

206. Reents H., Munch R., Dammeyer T., Jahn D., and Hartig E. (2006). The Fnr regulon of *Bacillus subtilis*. J. Bacteriol. *188*: 1103-1112.

207. Reizer J., Bachem S., Reizer A., Arnaud M., Saier M.H., Jr., and Stulke J. (1999). Novel phosphotransferase system genes revealed by genome analysis - the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*. Microbiology *145 (Pt 12)*: 3419-3429.

208. Ren D., Bedzyk L.A., Setlow P., England D.F., Kjelleberg S., Thomas S.M., Ye R.W., and Wood T.K. (2004). Differential gene expression to investigate the effect of (5Z)-4-bromo- 5-(bromomethylene)-3-butyl-2(5H)-furanone on *Bacillus subtilis*. Appl. Environ. Microbiol. *70*: 4941-4949.

209. Rice T.K., Schork N.J., and Rao D.C. (2008). Methods for handling multiple testing. Adv. Genet. *60*: 293-308.

210. Rockett J.C. and Hellmann G.M. (2004). Confirming microarray data - is it really necessary? Genomics *83*: 541-549.

211. Rouillard J.M., Zuker M., and Gulari E. (2003). OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. *31*: 3057-3062.

212. Rowland B.M., Grossman T.H., Osburne M.S., and Taber H.W. (1996). Sequence and genetic organization of a *Bacillus subtilis* operon encoding 2,3-dihydroxybenzoate biosynthetic enzymes. Gene *178*: 119-123.

213. Russell JB (1992). Another explanation for the toxicity of fermentation acids at low pH: Anion accumulation versus uncoupling. J. Appl. Microbiol. *73*: 363-370.

214. Sakamoto T. and Murata N. (2002). Regulation of the desaturation of fatty acids and its role in tolerance to cold and salt stress. Curr. Opin. Microbiol. *5*: 208-210.

215. Salzberg L.I. and Helmann J.D. (2007). An antibiotic-inducible cell wall-associated protein that protects *Bacillus subtilis* from autolysis. J. Bacteriol. *189*: 4671-4680.

216. Sarvas M., Harwood C.R., Bron S., and van Dijl J.M. (2004). Post-translocational folding of secretory proteins in Gram-positive bacteria. Biochim. Biophys. Acta *1694*: 311-327.

217. Schena M., Shalon D., Davis R.W., and Brown P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*: 467-470.

218. Schonert S., Seitz S., Krafft H., Feuerbaum E.A., Andernach I., Witz G., and Dahl M.K. (2006). Maltose and Maltodextrin Utilization by *Bacillus subtilis*. J. Bacteriol. *188*: 3911-3922.

219. Schumacher M.A., Allen G.S., Diel M., Seidel G., Hillen W., and Brennan R.G. (2004). Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. Cell *118*: 731-741.

220. Schumacher M.A., Seidel G., Hillen W., and Brennan R.G. (2006). Phosphoprotein Crh-Ser46-P displays altered binding to CcpA to effect carbon catabolite regulation. J. Biol. Chem. *281*: 6793-6800.

221. Schweder T. and Hecker M. (2004). Monitoring of stress responses. Adv. Biochem. Eng Biotechnol. *89*: 47-71.

222. Sebert M.E., Patel K.P., Plotnick M., and Weiser J.N. (2005). Pneumococcal HtrA protease mediates inhibition of competence by the CiaRH two-component signaling system. J. Bacteriol. *187*: 3969-3979.

223. Seidel G., Diel M., Fuchsbauer N., and Hillen W. (2005). Quantitative interdependence of coeffectors, CcpA and *cre* in carbon catabolite regulation of *Bacillus subtilis*. FEBS J. *272*: 2566-2577.

224. Serizawa M., Yamamoto H., Yamaguchi H., Fujita Y., Kobayashi K., Ogasawara N., and Sekiguchi J. (2004). Systematic analysis of SigD-regulated genes in *Bacillus subtilis* by DNA microarray and Northern blotting analyses. Gene *329*: 125-136.

225. Shivers R.P. and Sonenshein A.L. (2005). *Bacillus subtilis ilvB* operon: an intersection of global regulons. Mol. Microbiol. *56*: 1549-1559.

226. Sierro N., Makita Y., de H.M., and Nakai K. (2008). DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. Nucleic Acids Res. *36*: D93-D96.

227. Smits W.K., Dubois J.Y., Bron S., van Dijl J.M., and Kuipers O.P. (2005). Tricksy business: transcriptome analysis reveals the involvement of thioredoxin A in redox homeostasis, oxidative stress, sulfur metabolism, and cellular differentiation in *Bacillus subtilis*. J. Bacteriol. *187*: 3921-3930.

228. Sonenshein A.L. (2007). Control of key metabolic intersections in *Bacillus subtilis*. Nat. Rev. Microbiol. *5*: 917-927.

229. Sorek R. and Cossart P. (2010). Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nat. Rev. Genet. *11*: 9-16.

230. Spizizen J. (1958). Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. U. S. A *44*: 1072-1078.

231. Sprehe M., Seidel G., Diel M., and Hillen W. (2007). CcpA mutants with differential activities in *Bacillus subtilis*. J. Mol. Microbiol. Biotechnol. *12*: 96-105.

232. Steil L., Hoffmann T., Budde I., Volker U., and Bremer E. (2003). Genome-wide transcriptional profiling analysis of adaptation of *Bacillus subtilis* to high salinity. J. Bacteriol. *185*: 6358-6370.

233. Steil L., Serrano M., Henriques A.O., and Volker U. (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. Microbiology *151*: 399-420.

234. Stevens M.J. (2008).Transcriptome response of *Lactobacillus plantarum* to global regulator deficiency, stress and other environmental conditions. PhD dissertation.

235. Stulke J. and Hillen W. (2000). Regulation of carbon catabolism in *Bacillus* species**.** Annu. Rev. Microbiol. *54***:** 849-880.

236. Sturn A., Quackenbush J., and Trajanoski Z. (2002). Genesis: cluster analysis of microarray data**.** Bioinformatics. *18***:** 207-208.

237. Tam l.T., Eymann C., Albrecht D., Sietmann R., Schauer F., Hecker M., and Antelmann H. (2006). Differential gene expression in response to phenol and catechol reveals different metabolic activities for the degradation of aromatic compounds in *Bacillus subtilis***.** Environ. Microbiol. *8***:** 1408-1427.

238. Tang F., Barbacioru C., Nordman E., Li B., Xu N., Bashkirov V.I., Lao K., and Surani M.A. (2010). RNA-Seq analysis to capture the transcriptome landscape of a single cell**.** Nat. Protoc. *5***:** 516-535.

239. Tannler S., Fischer E., Le C.D., Doan T., Jamet E., Sauer U., and Aymerich S. (2008). CcpN controls central carbon fluxes in *Bacillus subtilis***.** J. Bacteriol. *190***:** 6178-6187.

240. Tatusov R.L., Natale D.A., Garkavtsev I.V., Tatusova T.A., Shankavaram U.T., Rao B.S., Kiryutin B., Galperin M.Y., Fedorova N.D., and Koonin E.V. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes**.** Nucleic Acids Res. *29***:** 22-28.

241. ter Beek A., Keijser B.J., Boorsma A., Zakrzewska A., Orij R., Smits G.J., and Brul S. (2008). Transcriptome analysis of sorbic acid-stressed *Bacillus subtilis* reveals a nutrient limitation response and indicates plasma membrane remodeling**.** J. Bacteriol. *190***:** 1751-1761.

242. Thackray P.D. and Moir A. (2003). SigM, an extracytoplasmic function sigma factor of *Bacillus subtilis*, is activated in response to cell wall antibiotics, ethanol, heat, acid, and superoxide stress**.** J. Bacteriol. *185***:** 3491-3498.

243. Thijs G., Marchal K., Lescot M., Rombauts S., De Moor B., Rouze P., and Moreau Y. (2002). A Gibbs sampling method to detect overrepresented motifs in the upstream regions of coexpressed genes**.** J. Comput. Biol. *9***:** 447-464.

244. Tojo S., Satomura T., Morisaki K., Deutscher J., Hirooka K., and Fujita Y. (2005). Elaborate transcription regulation of the *Bacillus subtilis ilv-leu* operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA**.** Mol. Microbiol. *56***:** 1560-1573.

245. Tojo S., Satomura T., Morisaki K., Yoshida K., Hirooka K., and Fujita Y. (2004). Negative transcriptional regulation of the *ilv-leu* operon for biosynthesis of branched-chain amino acids through the *Bacillus subtilis* global regulator TnrA**.** J. Bacteriol. *186***:** 7971-7979.

246. Tseng T.T., Gratwick K.S., Kollman J., Park D., Nies D.H., Goffeau A., and Saier M.H., Jr. (1999). The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins**.** J. Mol. Microbiol. Biotechnol. *1***:** 107-125.

247. Turinsky A.J., Grundy F.J., Kim J.H., Chambliss G.H., and Henkin T.M. (1998). Transcriptional activation of the *Bacillus subtilis ackA* gene requires sequences upstream of the promoter**.** J. Bacteriol. *180***:** 5961-5967.

248. Turner M.S. and Helmann J.D. (2000). Mutations in multidrug efflux homologs, sugar isomerases, and antimicrobial biosynthesis genes differentially elevate activity of the sigma(X) and sigma(W) factors in *Bacillus subtilis***.** J. Bacteriol. *182***:** 5202-5210.

249. van der Voort M. (2008). Fine tuning of the *Bacillus cereu*s stress response: role of transcriptional regulators. PhD dissertation.

250. van der Voort M., Kuipers O.P., Buist G., de Vos W.M., and Abee T. (2008). Assessment of CcpA-mediated catabolite control of gene expression in *Bacillus cereus* ATCC 14579**.** BMC. Microbiol. *8***:** 62.

251. van Hijum S.A., de Jong A., Baerends R.J., Karsens H.A., Kramer N.E., Larsen R., den Hengst C.D., Albers C.J., Kok J., and Kuipers O.P. (2005). A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data**.** BMC. Genomics *6***:** 77.

252. van Hijum S.A., de Jong A., Buist G., Kok J., and Kuipers O.P. (2003). UniFrag and GenomePrimer: selection of primers for genome-wide production of unique amplicons. Bioinformatics. *19*: 1580-1582.

253. van Hijum S.A., Garcia de la N.J., Trelles O., Kok J., and Kuipers O.P. (2003). MicroPreP: a cDNA microarray data pre-processing framework. Appl. Bioinformatics. *2*: 241-244.

254. van Hijum S.A., Medema M.H., and Kuipers O.P. (2009). Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. Microbiol. Mol. Biol. Rev. *73*: 481-509, Table.

255. van Vliet A.H. (2010). Next generation sequencing of microbial transcriptomes: challenges and opportunities. FEMS Microbiol. Lett. *302*: 1-7.

256. Veening J.W., Hamoen L.W., and Kuipers O.P. (2005). Phosphatases modulate the bistable sporulation gene expression pattern in *Bacillus subtilis*. Mol. Microbiol. *56*: 1481-1494.

257. Veening J.W., Smits W.K., and Kuipers O.P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. *62*: 193-210.

258. Wacker I., Ludwig H., Reif I., Blencke H.M., Detsch C., and Stulke J. (2003). The regulatory link between carbon and nitrogen metabolism in *Bacillus subtilis*: regulation of the *gltAB* operon by the catabolite control protein CcpA. Microbiology *149*: 3001-3009.

259. Wang Z., Gerstein M., and Snyder M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*: 57-63.

260. Wang Z.Q., Lawson R.J., Buddha M.R., Wei C.C., Crane B.R., Munro A.W., and Stuehr D.J. (2007). Bacterial flavodoxins support nitric oxide production by *Bacillus subtilis* nitric-oxide synthase. J. Biol. Chem. *282*: 2196-2202.

261. Warner J.B. and Lolkema J.S. (2003). CcpA-dependent carbon catabolite repression in bacteria. Microbiol. Mol. Biol. Rev. *67*: 475-490.

262. Waters L.S. and Storz G. (2009). Regulatory RNAs in bacteria. Cell *136*: 615-628.

263. Weber M.H., Klein W., Muller L., Niess U.M., and Marahiel M.A. (2001). Role of the *Bacillus subtilis* fatty acid desaturase in membrane adaptation during cold shock. Mol. Microbiol. *39*: 1321-1329.

264. Wecke T., Zuhlke D., Mader U., Jordan S., Voigt B., Pelzer S., Labischinski H., Homuth G., Hecker M., and Mascher T. (2009). Daptomycin versus Friulimicin B: in-depth profiling of *Bacillus subtilis* cell envelope stress responses. Antimicrob. Agents Chemother. *53*: 1619-1623.

265. Weickert M.J. and Adhya S. (1992). A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. *267*: 15869-15874.

266. Weickert M.J. and Chambliss G.H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A *87*: 6238-6242.

267. Westers H., Westers L., Darmon E., van Dijl J.M., Quax W.J., and Zanen G. (2006). The CssRS two-component regulatory system controls a general secretion stress response in *Bacillus subtilis*. FEBS J. *273*: 3816-3827.

268. Westers L., Westers H., and Quax W.J. (2004). *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. Biochim. Biophys. Acta *1694*: 299-310.

269. Wiegert T., Homuth G., Versteeg S., and Schumann W. (2001). Alkaline shock induces the *Bacillus subtilis* sigma(W) regulon. Mol. Microbiol. *41*: 59-71.

270. Wilks J.C., Kitko R.D., Cleeton S.H., Lee G.E., Ugwu C.S., Jones B.D., BonDurant S.S., and Slonczewski J.L. (2009). Acid and base stress and transcriptomic responses in *Bacillus subtilis*. Appl. Environ. Microbiol. *75*: 981-990.

271. Winstedt L. and von W.C. (2000). Terminal oxidases of *Bacillus subtilis* strain 168: one quinol oxidase, cytochrome aa(3) or cytochrome bd, is required for aerobic growth. J. Bacteriol. *182*: 6557-6564.

272. Yamamoto H., Serizawa M., Thompson J., and Sekiguchi J. (2001). Regulation of the glv operon in *Bacillus subtilis*: YfiA (GlvR) is a positive regulator of the operon that is repressed through CcpA and *cre*. J. Bacteriol. *183*: 5110-5121.

273. Ye R.W., Tao W., Bedzyk L., Young T., Chen M., and Li L. (2000). Global gene expression profiles of *Bacillus subtilis* grown under anaerobic conditions**.** J. Bacteriol. *182***:** 4458-4465.

274. Ye R.W., Wang T., Bedzyk L., and Croker K.M. (2001). Applications of DNA microarrays in microbial systems**.** J. Microbiol. Methods *47***:** 257-272.

275. Yeats C., Rawlings N.D., and Bateman A. (2004). The PepSY domain: a regulator of peptidase activity in the microbial environment? Trends Biochem. Sci. *29***:** 169-172.

276. Yoshida K., Kobayashi K., Miwa Y., Kang C.M., Matsunaga M., Yamaguchi H., Tojo S., Yamamoto M., Nishi R., Ogasawara N. et al. (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis***.** Nucleic Acids Res. *29***:** 683-692.

277. Yoshida K., Ohki Y.H., Murata M., Kinehara M., Matsuoka H., Satomura T., Ohki R., Kumano M., Yamane K., and Fujita Y. (2004). *Bacillus subtilis* LmrA is a repressor of the *lmrAB* and *yxaGH* operons: identification of its binding site and functional analysis of *lmrB* and *yxaGH***.** J. Bacteriol. *186***:** 5640-5648.

278. Yoshida K., Yamaguchi H., Kinehara M., Ohki Y.H., Nakaura Y., and Fujita Y. (2003). Identification of additional TnrA-regulated genes of *Bacillus subtilis* associated with a TnrA box**.** Mol. Microbiol. *49***:** 157-165.

279. Yoshida K.I., Fujita Y., and Ehrlich S.D. (2000). An operon for a putative ATP-binding cassette transport system involved in acetoin utilization of *Bacillus subtilis***.** J. Bacteriol. *182***:** 5454-5461.

280. Yuen T., Wurmbach E., Pfeffer R.L., Ebersole B.J., and Sealfon S.C. (2002). Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays**.** Nucleic Acids Res. *30***:** e48.

281. Zalieckas J.M., Wray L.V., Jr., Ferson A.E., and Fisher S.H. (1998). Transcription-repair coupling factor is involved in carbon catabolite repression of the *Bacillus subtilis* hut and gnt operons**.** Mol. Microbiol. *27***:** 1031-1038.

282. Zalieckas J.M., Wray L.V., Jr., and Fisher S.H. (1998). Expression of the *Bacillus subtilis acsA* gene: position and sequence context affect *cre*-mediated carbon catabolite repression**.** J. Bacteriol. *180***:** 6649-6654.

283. Zanen G., Houben E.N., Meima R., Tjalsma H., Jongbloed J.D., Westers H., Oudega B., Luirink J., van Dijl J.M., and Quax W.J. (2005). Signal peptide hydrophobicity is critical for early stages in protein export by *Bacillus subtilis***.** FEBS J. *272***:** 4617-4630.

284. Zellmeier S., Hofmann C., Thomas S., Wiegert T., and Schumann W. (2005). Identification of sigma(V)-dependent genes of *Bacillus subtilis***.** FEMS Microbiol. Lett. *253***:** 221-229.

285. Zhang S. and Haldenwang W.G. (2003). RelA is a component of the nutritional stress activation pathway of the *Bacillus subtilis* transcription factor sigma B**.** J. Bacteriol. *185***:** 5714-5721.

286. Zhang W., Li F., and Nie L. (2010). Integrating multiple 'omics' analysis for microbial biology: application and methodologies**.** Microbiology *156***:** 287-301.

287. Zheng G., Hehn R., and Zuber P. (2000). Mutational analysis of the *sbo-alb* locus of *Bacillus subtilis*: identification of genes required for subtilosin production and immunity**.** J. Bacteriol. *182***:** 3266-3273.

288. Zheng G., Yan L.Z., Vederas J.C., and Zuber P. (1999). Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilosin**.** J. Bacteriol. *181***:** 7346-7355.

289. Zomer A.L. (2007). Adaptation of *Lactococcus lactis* to its environment. PhD dissertation.

290. Zomer A.L., Buist G., Larsen R., Kok J., and Kuipers O.P. (2007). Time-resolved determination of the CcpA regulon of *Lactococcus lactis subsp. cremoris* MG1363**.** J. Bacteriol. *189***:** 1366-1381.

Afterword & acknowledgements

"*Being a graduate student is like becoming all of the Seven Dwarves. In the beginning you're Dopey and Bashful. In the middle, you are usually sick (Sneezy), tired (Sleepy), and irritable (Grumpy). But at the end, they call you Doc, and then you're Happy*."
(Ronald T. Azuma, "So long and thanks for the Ph.D.").

I cannot recall anymore when I came across this quotation, but I do remember I was kind of desperate and felt like "Grumpy" back then. In the middle of 2004 having no publishable results, I could hardly believe that the "Doc & Happy" stage of my PhD time would ever become a reality. For one or another reason, during a 9-month prolongation of my contract I started to foster an attitude of "no matter what and when but I will make it".
It was a long and bumpy journey, which required a lot of tenacity and self-discipline, both lacking every now and then, but against all odds I managed to reach the "Doc" phase. When time progressed, I realized that next to the scientific achievements, also other objectives should be reached during the PhD pathway. What mattered a lot to me was that during this period at the University of Groningen I developed various skills and personality traits that served as an excellent basis for my future professional career. Most importantly, I also learned that *success in life comes not from having the right cards, but from playing bad ones properly* (Joshua Dool).

There are many people who deserve a special mention for their support during this whole adventure between 2002-2007 in Haren and later on in Nijmegen and Boxmeer.
First of all, Oscar, I would like to thank you for offering me the opportunity to work on this challenging project. The project probably did not go precisely as we both would have wished and so I do appreciate that you showed your skills as promotor through a bad patch and helped me a great deal on my way to the "Doc" phase.
Girbe, "captain my captain", you were always busy with many projects and your teaching courses, still you were always eager to help. We were both struggling to find the golden treasure in the huge data sets of the CcpA time series and I learned a lot from your comments especially while writing my first publication as well as preparing the rest of my thesis. Jan, thank you for your contribution in the initial phase of the IOP project.
Aldert, I really appreciated your help with computer and motif searching tasks and that you shared your optimism with me. Jolanda and Henk, it was great to have you as undergraduate students. Jolanda, despite having to redo all the CcpA work on the new generation oligo microarrays, you did not get discouraged and decided to start your own

PhD project as well. Henk, thanks for your work (especially all the growth curves) on the lactic acid stress project.
My office mates, Wiep Klaas and Jan-Willem, it was a "peculiar" experience to share the tiny office with you: weird "noises", unusual type of "decorative" wallpaper (i.e. the first pages of your publications) and other habits, which I could keep on listing …. . Somehow, I managed to keep sane and looking back on it after years I must seriously admit it was kind of funny. Thanks for your input in the secretion stress paper and valuable discussions about other projects of mine. I strongly believe that both of you will reach the top in science! Jan-Willem, I appreciated that you had tried to join me in the x^{th} PCR amplification round. You had a chance to experience with what kind of task I was dealing with and I got an unforgettable lesson on a very unusual cursing in Dutch. Wiep Klaas, let me pass the *B. subtilis* microarray work over in silence ;) … I am grateful not only for all the discussions with you on a wide variety of subjects but also for the fact that you shared your enthusiasm for research (for example with the *yjbf* project).
Asia, we shared lots of ups and downs, but finally we made it to this stage! Evert-Jan, I could speak volumes about you and our collaboration on the big time series project, which hopefully will get published after all the time and effort we invested in it! You managed to get my adrenaline level to the maximum limit and to be honest I am glad I am still alive after the fermentor explosions we experienced. Thanks for the FIVA!
Siger, we went through thick and thin together and never gave up on the PCR amplification or the microarray production/optimization. Harma, thanks for the contribution in the lactic acid stress chapter. Sacha, without you it would be impossible to harness the microarray data! Although I spent most of my time in the array lab, I would like to thank my lab mates: Reindert, Akos, Ola, Mariska and Elise for the fun we shared. Łukasz, the coffee breaks with you were refreshing! I could keep on listing the names of the current and the past members of the MolGen group, but I would simply like to express my sincere gratitude to all of you for the unforgettable atmosphere!
Adrian and Ola, it is great to have you as friends!

Furthermore I would like to mention my superiors at Intervet/Schering-Plough Animal Health, Marcus and Marc (and in the past Hans and Emiel), for your empathy and giving me the space and time I needed to complete this thesis. Dominique, thank you for helping me out with my everyday Dutch; sharing the office with you was a pleasure. Klaas and Wiep Klaas thanks for reviewing the Dutch summary. Peter, I am grateful for your valuable language corrections. The whole Regulatory Affairs department, thanks for the gezelligheid.

Asia, Andrzej, Iza and Piotr thanks for your friendliness and a warm welcome you gave us in Nijmegen.

Moim rodzicom i teściom, dziękuje za wszystko!!!

Finally, my wife and children. Kubuś i Adaś, you brought incredible joy and another dimension in my life. Now, I know for sure that *children seldom misquote, in fact, they usually repeat word for word what you shouldn't have said*. Wielke przytulańce i buziaki od taty! Martuś, there are actually no words to commit my feelings to paper. You helped me a lot by arranging all the practical issues for the defence and getting the book ready for printing. It was great to work together on the Polish summary (buźka za rysunki!) and the cover. Thank you for your patience, open heart and unwavering faith in me. If it had not been for your support and unconditional love, I would not have reached this place. *Loveu*!!!

Nothing in the world can take the place of persistence. Talent will not ... genius will not ... education will not ... persistence and determination are omnipotent.
(President Calvin Coolidge)